

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:38:47 ; Search time 8644.89 Seconds
(without alignments)
12117.322 Million cell updates/sec

Title: US-10-624-932-1
Perfect score: 2752
Sequence: 1 ccgcggggcccgcgcccg.....tgagtgcctgaggccggccag 2752

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	960.2	34.9	2802	9	AY406491	AY406491 Homo sapi
2	950.4	34.5	2791	9	AY406493	AY406493 Mus muscu
3	923.4	33.6	3790	3	AK031655	AK031655 Mus muscu
4	884	32.1	1852	3	CR598115	CR598115 full-leng
5	872.4	31.7	3866	3	AK018177	AK018177 Mus muscu
6	814	29.6	2802	9	AY406492	AY406492 Pan trogl
7	810.4	29.4	2532	9	AY411747	AY411747 Homo sapi
8	780.4	28.4	2532	9	AY411749	AY411749 Mus muscu

	9	768.6	27.9	1034	4	BI758231	BI758231	603029876
	10	736.6	26.8	1532	3	BC033727	BC033727	Homo sapi
	11	735.2	26.7	788	1	AI951556	AI951556	wv36f04.x
c	12	721.6	26.2	796	5	BX348193	BX348193	BX348193
	13	716.4	26.0	818	4	BI818609	BI818609	603033362
	14	692	25.1	853	5	BX364574	BX364574	BX364574
	15	678.2	24.6	2532	9	AY411748	AY411748	Pan trogl
	16	672.6	24.4	934	2	BF311804	BF311804	601897316
	17	666	24.2	900	5	BX345406	BX345406	BX345406
	18	665	24.2	859	2	BF311896	BF311896	601897733
	19	645.2	23.4	977	5	BX345407	BX345407	BX345407
	20	631.2	22.9	1072	5	BX422753	BX422753	BX422753
	21	618.8	22.5	756	5	BU612387	BU612387	UI-M-EW0-
	22	613.2	22.3	874	5	BQ689148	BQ689148	AGENCOURT
	23	610.2	22.2	2775	9	AY401471	AY401471	Mus muscu
	24	607	22.1	889	5	BQ691915	BQ691915	AGENCOURT
	25	599	21.8	2775	9	AY401469	AY401469	Homo sapi
	26	595.2	21.6	601	1	AL516580	AL516580	AL516580
	27	589.4	21.4	604	2	BE314370	BE314370	601147261
	28	566.6	20.6	678	6	CA749784	CA749784	UI-M-FD0-
	29	561	20.4	1175	2	BF530640	BF530640	602071931
	30	555.6	20.2	788	6	CA317532	CA317532	UI-M-FW0-
	31	555	20.2	572	7	CR554569	CR554569	DKFZp459I
	32	494.2	18.0	499	7	CR747398	CR747398	CR747398
	33	486.2	17.7	548	5	BX452510	BX452510	BX452510
	34	475.4	17.3	824	4	BI737024	BI737024	603360874
	35	471.2	17.1	540	4	BI467799	BI467799	389583 MA
	36	470.8	17.1	2507	9	AY401470	AY401470	Pan trogl
	37	438.2	15.9	675	6	CA315487	CA315487	UI-M-FW0-
	38	433	15.7	471	5	BX282095	BX282095	BX282095
	39	428.4	15.6	1147	5	BU840446	BU840446	AGENCOURT
	40	424	15.4	843	4	BG913440	BG913440	602811321
	41	417.8	15.2	856	7	CN164143	CN164143	994266 MA
	42	411.8	15.0	749	7	CF735417	CF735417	UI-M-HB0-
	43	410.6	14.9	1111	4	BG298307	BG298307	602397080
	44	409.8	14.9	751	7	CF735550	CF735550	UI-M-HB0-
	45	408.8	14.9	460	2	BF443156	BF443156	260542 MA

ALIGNMENTS

RESULT 1

AY406491

LOCUS AY406491 2802 bp DNA linear GSS 15-DEC-2003

DEFINITION Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY406491

VERSION AY406491.1 GI:39762465

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2802)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 2802)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES Location/Qualifiers
 source 1. .2802
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <1. .>2802
 /gene="UNC5C"
 /locus_tag="HCM2575"
 ORIGIN

Query Match 34.9%; Score 960.2; DB 9; Length 2802;
 Best Local Similarity 61.3%; Pred. No. 6.4e-192;
 Matches 1682; Conservative 0; Mismatches 995; Indels 69; Gaps 6;

Qy	57	CCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCTCGCCGCTTGGCTCCGCGGCTCGGG	116
Db	66	GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA	125
Qy	117	TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC	176
Db	126	TGATGACTTTTTTTCATGAACTCCAGAACTTTTCTCTGATCCACCTGAGCCTCTGCC	185
Qy	177	CCACTTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT	236
Db	186	ACATTTCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA	245
Qy	237	GTGCAAGGCCGTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCA	296
Db	246	CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCA	305
Qy	297	GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGT	356
Db	306	GAAGGACCACATAGTAGATGAAAGAGTAGATGAACTTCCGGTCTCATTGTCCGGGAAGT	365
Qy	357	CCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGCTGGAGGAATACTGGTG	416
Db	366	GAGCATTGAGATTTTCGCCCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTG	425
Qy	417	CCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCAT	476
Db	426	CCAGTGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCAT	485

Qy		477 AGC-----CAGATTGCGCAAGAACTTCGAGCAGGACGGCTGGCCAAGGAGGTGTCCTT	530
Db		486 TGCATNNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGCTTTT	545
Qy		531 GGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA	590
Db		546 GGAACAGGAAGTCTTACTCCAGTGTGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGA	605
Qy		591 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAACATCACGCG	650
Db		606 ATGGTTGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT	665
Qy		651 GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT	710
Db		666 TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGT	725
Qy		711 GGCCAAGAACATCGTGGCACGTGCCCCGAGCGCCTCCGCTGCTGTTCATCGTCTACGTGAA	770
Db		726 TGCCAAAACATTGTTGCCAAGAGGAAAAGTACAACCTGCCACTGTCATAGTCTATGTCAA	785
Qy		771 CGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG	830
Db		786 CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTA	845
Qy		831 GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGGCGCTTTCTGTGA	890
Db		846 TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCCTCAATGGGGGTGCCTTCTGTGA	905
Qy		891 GGGGCAGAATGTCCAGAAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAG	950
Db		906 AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGAC	965
Qy		951 CCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTG	1010
Db		966 GCCATGGAGCAAGTGGTCTACTTGTGGAAGTGAAGTGCACCCACTGGCGCAGGAGGGAGTG	1025
Qy		1011 CTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCG	1070
Db		1026 CACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAA	1085
Qy		1071 CAACTGTACCACTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTA	1130
Db		1086 GAAGTGCAGTGTGAGGCTTTGCATGCAGACTGCTCCTGATTGAGATGATGTTGCTCTCTA	1145
Qy		1131 TGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCTCTGCTGCTGCTGCTCCTCATCCT	1187
Db		1146 TGTGAGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTT	1205
Qy		1188 CGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC	1247
Db		1206 TGTGTATCGGAAGAATCATCGTGAAGTGTGAGTCAAGATATTGACTCTTCGGCACTCAA	1265
Qy		1248 CTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCAC	1307
Db		1266 TGGGGGCTTTTCAGCCTGTGAACATCAAG-----GCAGCAAGACAAGATCTGCTGGC	1316

Qy	1308	CATCCAGCCGGACCTCAGCACACCACCACCACCTACCAGGGCAGTCTCTGTCCCCCGGCA	1367
Db	1317	TGTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCA	1376
Qy	1368	GGATG-----GGCCCAGCCCCAAGTTCCAGCTCACCAAT----GGGCACCTGCTCAGCCC	1418
Db	1377	TGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCT	1436
Qy	1419	CCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTT	1478
Db	1437	GAAAATCAAAGTGTACAACACCTCAGGTGCTGTACCCCCCAAGATGACCTCTCTGAGTT	1496
Qy	1479	CGTCTCCCGCCTCTCCACCCAGAATACTTCCGCTCCCTGCCCGGAGGCACCAGCAACAT	1538
Db	1497	TACGTCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCT	1556
Qy	1539	GA-----CCTATGGGACCTTCAA	1556
Db	1557	GAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAA	1616
Qy	1557	CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA	1616
Db	1617	CTCGCTGGGAGGTACCTTATTGTTCCCAATTGAGGAGTCAGCTTGCTGATTCCCGCTGG	1676
Qy	1617	TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT	1676
Db	1677	GGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAACTAT	1736
Qy	1677	GAGGTTGCCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC	1736
Db	1737	GAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCC	1796
Qy	1737	TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC	1796
Db	1797	AGGAGCTCTGCTCACCCGCCCAGTCGTCCTCACTATGCATCACTGCGCAGACCCCAATAC	1856
Qy	1797	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	1856
Db	1857	CGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAGGATGTGGT	1916
Qy	1857	GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	1916
Db	1917	GGTGGTCGGGGAGGAAAACCTTACCACCCCCTGCTACATTGAGCTGGATGCAGAGGCCCTG	1976
Qy	1917	CTACGTCTTCACCGAGCAGCTGGGCGCCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	1976
Db	1977	CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACCAAAGC	2036
Qy	1977	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA	2036
Db	2037	GGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTGGAGTA	2096
Qy	2037	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	2096
Db	2097	CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT	2156
Qy	2097	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAG	2156

```

      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2157 TGAGAGACAGATGGGAGGACAGCTCCTAGAGAAGCTAAGGCTCTTCATTTTAAAGGCAG 2216
Qy      2157 TTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT 2216
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2217 CACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCTCTGGAAGAGCAAATT 2276
Qy      2217 CCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAAGT 2276
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2277 GCTGGCTAAATATCAGGAAATTCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCT 2336
Qy      2277 GCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2336
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2337 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACCT 2396
Qy      2337 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2396
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2397 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGA 2456
Qy      2397 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTCCCAGCCCTGGT 2456
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2457 GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACACGGTTCAC 2516
Qy      2457 GGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTGCGCAGAAGATAATTTCCAGCCTGGA 2516
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2517 GGGGCCAGTGCTTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576
Qy      2517 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAG 2576
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2577 TGCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCATAAGCTGAACCTGGACAG 2636
Qy      2577 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2636
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2637 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696
Qy      2637 GGCGCGGCACCTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2697 AGCACAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGG 2756
Qy      2697 CCAGCCAGACGCTGGCCTCTTACAGTGTGCGAGGCTGAGTGCTGA 2742
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2757 AAGACATGAAACGGTGGTGTCTTAGCAGCAGAAGGGCAGTATTAA 2802

```

RESULT 2

AY406493

LOCUS AY406493 2791 bp DNA linear GSS 15-DEC-2003

DEFINITION Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY406493

VERSION AY406493.1 GI:39762467

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2791)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2791)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1. .2791

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

gene <1..>2791

/gene="UNC5C"

/locus_tag="HCM2575"

ORIGIN

Query Match 34.5%; Score 950.4; DB 9; Length 2791;

Best Local Similarity 62.3%; Pred. No. 7.5e-190;

Matches 1653; Conservative 0; Mismatches 921; Indels 80; Gaps 7;

Qy 149 TGCCTGGTGCCAAACCCGACCTGCTTCCCCACTTCCTGGTGGAGCCCCGAGGATGTGTACA 208

| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 158 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCCGAGGAAGCTTACA 217

Qy 209 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCCGCCACGCAGATCTTCT 268

| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 218 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 277

Qy 269 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 328

| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 278 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 337

Qy 329 GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGA 388

| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 338 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTACGCCAGCAGGTGGAGG 397

Qy 389 AGGTGTTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 448

| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 398 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 457

Qy 449 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAG-----ATTGCGCAAGAACTTCGAGC 502

| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 458 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTTTCCAGATCTGCGGAAGACATTTCGAGC 517

Qy 503 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 562

Db	518	AGGAACCCTTGGGAAAGGAAGTGTCTTGGAGCAGGAAGTCTTACTCCAGTGTGCGCCAC	577
Qy	563	CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT	622
Db	578	CTGAAGGGATCC-----GTAGAATGGCTAAAGAATGAAGACATAATTGATCCTG	626
Qy	623	CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCC	682
Db	627	CTGAAGATCGGAACCTTTTATATTACTATCGATCACAACTGATCATCAAGCAAGCCCCGAC	686
Qy	683	TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTGCGCCGACGG	742
Db	687	TCTCAGATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAGCA	746
Qy	743	CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCG	802
Db	747	CCACAGCCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTG	806
Qy	803	TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG	862
Db	807	TGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAG	866
Qy	863	CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCA	922
Db	867	CCCCACTCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTA	926
Qy	923	CCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTTCGGCCTGTGGGCTGG	982
Db	927	CGTTATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTG	986
Qy	983	ACTGCACCCACTGGCGGAGCCGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGGAGG	1042
Db	987	AATGCACCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCAAGAACGGGGGTAAGG	1046
Qy	1043	AGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTG	1102
Db	1047	ACTGTGATGGCCTGGTCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCAGGCTG	1106
Qy	1103	CTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCC	1159
Db	1107	CTCCTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTC	1166
Qy	1160	TGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT	1219
Db	1167	TGGCGATCACTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGT	1226
Qy	1220	CAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCA	1279
Db	1227	CTGACATCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAG----	1282
Qy	1280	GCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCA	1339
Db	1283	-----GCTGCCAGACAAGATCTCCTGGCTGTCCCCCTGACCTCACCTCAGCTGCAGCCA	1337
Qy	1340	CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATG-----GGCCCAGCCCCAAGTTCCAGCT	1394

Db 1338 TGTACAGGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACT 1397

Qy 1395 CACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTC 1454
| | | | | | | | | | | | | | | | | | | | | |

Db 1398 CTCCAATTCTGGACCCACTACCCAACCTGAAAATCAAAGTGTACAACAGCTCAGGTGCTG 1457

Qy 1455 TCCCACCTC----TGAGGCCGAGGAGTTTCGTCTCCCGCCTCTCCACCC----- 1498
| | | | | | | | | | | | | | | | | | | | | |

Db 1458 TCACTCCTCAGGATGACCTTGCCGAGTTCTCATCCAACTGTCACCCAGATGACCCAGT 1517

Qy 1499 -----AGAACTACTTCCGCTCCCTGCCCCGAGGCA 1528
| | | | | | | | | | | | | | | | | | | | | |

Db 1518 CCTTGCTAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGACTGACC 1577

Qy 1529 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1588
| | | | | | | | | | | | | | | | | | | | | |

Db 1578 CATCCTGCACAGCATTGTTGGTACCTTCAACTCTCTTGGGGGTACCTCATCATTCCTAATT 1637

Qy 1589 CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1648
| | | | | | | | | | | | | | | | | | | | | |

Db 1638 CAGGAGTAAGCTTGCTGATTCCCGCTGGGGCCATTCTCAGGGGAGAGTCTATGAAATGT 1697

Qy 1649 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGC 1708
| | | | | | | | | | | | | | | | | | | | | |

Db 1698 ATGTGACTGTACACAGGAAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTAC 1757

Qy 1709 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG 1768
| | | | | | | | | | | | | | | | | | | | | |

Db 1758 TTACCCCTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTATCCTCA 1817

Qy 1769 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1828
| | | | | | | | | | | | | | | | | | | | | |

Db 1818 CTCTGCATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGG 1877

Qy 1829 CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCT 1888
| | | | | | | | | | | | | | | | | | | | | |

Db 1878 CAGTGCAGGGACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCT 1937

Qy 1889 ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG 1948
| | | | | | | | | | | | | | | | | | | | | |

Db 1938 GTTACATTAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATG 1997

Qy 1949 CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 2008
| | | | | | | | | | | | | | | | | | | | | |

Db 1998 CCCTGGTTGGGCAGTCCACCACCAAAGCAGTGCCAAGCGTCTTAAACTGGCCATCTTTG 2057

Qy 2009 CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCC 2068
| | | | | | | | | | | | | | | | | | | | | |

Db 2058 GGCCCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACAC 2117

Qy 2069 ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGG 2128
| | | | | | | | | | | | | | | | | | | | | |

Db 2118 AGGATGCCCTGAAGGAAGTTCTACAACCTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAG 2177

Qy 2129 AGCCACGGGTCTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATG 2188
| | | | | | | | | | | | | | | | | | | | | |

Db 2178 AACCCAAGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTTCATGACA 2237

Qy 2189 TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATC 2248
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2238 TCGCCCATTCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCATTTTACC 2297

Qy 2249 ACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCC 2308
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2298 ACATCTGGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCC 2357

Qy 2309 CCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGA 2368
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2358 TAAACACAGTGGAAGTGGTTTGCAAACCTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGA 2417

Qy 2369 GCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGG 2428
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2418 TCTTCCAGCTCAACTGTACTGTGTGAGAGGAACCTACTGGCATCGACTTACCTCTCCTGG 2477

Qy 2429 AGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCA 2488
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2478 ACCCTGCTAGTACCATCACCCTGTACCGGACCAAGTGCTTTCAGCATTCTCTCCCTA 2537

Qy 2489 TTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGA 2548
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2538 TCCGGCAGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGA 2597

Qy 2549 CTCTGGCCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCC 2608
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2598 TGCTGGCCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCC 2657

Qy 2609 CCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCC 2668
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2658 CAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCA 2717

Qy 2669 AGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGG 2728
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2718 TGCTGGCAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCTTGGCAGCAG 2777

Qy 2729 AGGCTGAGTGCTGA 2742
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2778 AAGGACAGTATTGA 2791

RESULT 3

AK031655

LOCUS AK031655 3790 bp mRNA linear HTC 03-APR-2004

DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030473H24 product:unc5 homolog (C. elegans) 3, full insert sequence.

ACCESSION AK031655

VERSION AK031655.1 GI:26327502

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 3790)
 AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
 Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details:
URL:<http://genome.gsc.riken.jp/>
URL:<http://fantom.gsc.riken.jp/>.

```

FEATURES
source
1. .3790
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030473H24"
/db_xref="taxon:10090"
/clone="6030473H24"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
CDS
118. .2970
/note="unnamed protein product; putative
unc5 homolog (C. elegans) 3 (MGD|MG1:1095412,
GB|NM_009472, evidence: BLASTN, 99%, match=464)"
/codon_start=1
/protein_id="BAC27495.1"
/db_xref="GI:26327503"
/translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE
FFHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPVNLVCKASPATQIYFKCNSEWVHQ
KDHYVVDVERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
RIAYLRKTTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
TIDHNLIIKQARLSDTANYTCAKNIVAKRKSTTATVIVVYNGGWSTWTEWSVCNSRC
GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTSWSKWSTCGTECTH
WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGGFIYPISTEHRPQNEYGFSSAPD
SDDVALYVGIVIAVTVCLAITVVVALFVYRKNNHRDFESDIIDSSALNGGFQPVNIKAA
RQDLLAVPPDLTSAAAMYRGPVYALHDVSDKI PMTNSPILDPLPNLKI KVNSSGAVT
PQDDLAEFSSKLS PQMTQS LLENEALNLKNQSLARQTDPSCTAFGT FNSLGGHLIIPN
SGVSL LI PAGAI PQGRVYEMYVT VHRKENMRPPMEDSQTLLTPV VSCGPPGALLTRPV
ILT LHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTT PCYIQLDAEACHILTEN
LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM
GGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSLLAKYQEIPFYHIWSGSQRNLHC
TFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEPTGIDLPLLDPASTITVT
GPSAISIPLRQKLCSSLDAPQTRGHDWRMLAHKLNLDRLYLN YFATKSSPTGVILDL
WEAONFPDGNLSMLAAVLEEMGRHETVVYLA AEGOY"

```

ORIGIN

Query Match 33.6%; Score 923.4; DB 3; Length 3790;
Best Local Similarity 61.3%; Pred. No. 3.9e-184;
Matches 1659; Conservative 0; Mismatches 926; Indels 120; Gaps 6;

Qy	149	TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCAGGATGTGTACA	208
Db	275	TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCAGGAAGCTTACA	334
Qy	209	TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT	268
Db	335	TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT	394

Qy 269 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 328
 ||||| ||||| | ||||| ||||| | || ||||| ||||| |||||
 Db 395 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 454

Qy 329 GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGA 388
 | | |||| | | |||| ||||| | |||| | ||||| |||||
 Db 455 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTACGCCAGCAGGTGGAGG 514

Qy 389 AGGTGTTCTGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 448
 | |||| |||| || || ||||| ||||| ||||| ||||| | ||||| |
 Db 515 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 574

Qy 449 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC 508
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 575 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTGAGCAGGAAC 634

Qy 509 CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG 568
 | |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 635 CCTTGGGAAAGGAAGTGTCTTGGAGCAGGAAGTCTTACTCCAGTGTGCGCCACCTGAAG 694

Qy 569 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG 628
 | ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 695 GGATCCCAGTGGCTGAGGTAGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG 754

Qy 629 ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG 688
 | | || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 755 ATCGGAACCTTTTATATTACTATCGATCACAACTGATCATCAAGCAAGCCCGACTCTCAG 814

Qy 689 ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTGCGCCGAGCGCCTCCG 748
 | |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 815 ATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAGCACCACAG 874

Qy 749 CTGCTGTCATCGTCTACGTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCA 808
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 875 CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTGTA 934

Qy 809 GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC 868
 | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 935 ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC 994

Qy 869 TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT 928
 |||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 995 TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTACGTTAT 1054

Qy 929 GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA 988
 | |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1055 GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAATGCA 1114

Qy 989 CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC 1048
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1115 CCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCAAGAACGGGGGTAAAGGACTGTG 1174

Qy 1049 AGGGCACTGACCTGGACACCCGCAACTGTACCACTGACCTCTGTGTACACAG----- 1100
 | |||| | |||| | |||| ||||| ||||| ||||| ||||| |||||
 Db 1175 ATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCAGGGATTTCATTT 1234

Qy 1101 -----TGCTTCTGGCC 1111
 Db 1235 ACCCCATTTCAACTGAGCACAGACCCAGAAATGAATATGGATTTTCTTCTGCTCCTGACT 1294
 Qy 1112 CTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCTCTGC 1168
 Db 1295 CAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCTGGCGATCA 1354
 Qy 1169 TGCTGCTTGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGG 1228
 Db 1355 CTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACATCA 1414
 Qy 1229 CTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAG 1288
 Db 1415 TTGACTCCTCAGCACTCAATGGCGGCTTTGAGCCTGTGAACATCAAG-----GCTG 1465
 Qy 1289 ACAACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGG 1348
 Db 1466 CCAGACAAGATCTCCTGGCTGTCCCCCCTGACCTCACCTCAGCTGCAGCCATGTACAGGG 1525
 Qy 1349 GCAGTCTCTGTCCCGGCGAGGATG-----GGCCAGCCCCAAGTTCCAGCTCACCAATGG 1403
 Db 1526 GACCTGTCTATGCTCTGCATGATGTCTCAGACAAATCCCAATGACCAACTCTCCAATTC 1585
 Qy 1404 GCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTC 1463
 Db 1586 TGGACCCACTACCCAACCTTGAAGATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTC 1645
 Qy 1464 ----TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC----- 1498
 Db 1646 AGGATGACCTTGCCGAGTTCTCATCAAACCTGTACCCAGATGACCCAGTCCCTTGCTAG 1705
 Qy 1499 -----AGAACTACTTCCGCTCCCTGCCCGGAGGCACCAGCAACA 1537
 Db 1706 AGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGACTGACCCATCCTGCA 1765
 Qy 1538 TGACCTATGGGACCTTCAACTTCCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCA 1597
 Db 1766 CAGCATTTGGTACCTTCAACTCTCTTGGGGGTACCTCATCATTCCTAATTCAGGAGTAA 1825
 Qy 1598 GCCTCCTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGC 1657
 Db 1826 GCTTGCTGATTCCCGCTGGGGCCATTCTCAGGGGAGAGTCTATGAAATGTATGTGACTG 1885
 Qy 1658 TGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCA 1717
 Db 1886 TACACAGGAAAGAAAATATGAGGCCCCCATGGAAGACTCTCAGACCCTACTTACCCCTG 1945
 Qy 1718 TCGTTAGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACC 1777
 Db 1946 TGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTATCCTCACTCTGCATC 2005
 Qy 1778 ACTGTGGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGG 1837
 Db 2006 ACTGTGCAGACCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACAGGCAGTGCAGG 2065
 Qy 1838 GCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCC 1897

Db	2066	GACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCCTGTTACATTC	2125
Qy	1898	AGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGG	1957
Db	2126	AGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGGTTG	2185
Qy	1958	GAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGG	2017
Db	2186	GGCAGTCCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCTCT	2245
Qy	2018	CCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCAC	2077
Db	2246	GCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATGCCC	2305
Qy	2078	TCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGG	2137
Db	2306	TGAAGGAAGTTCTACAACCTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGG	2365
Qy	2138	TCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCT	2197
Db	2366	CTCTTCGTTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTATGACATCGCCCATT	2425
Qy	2198	CCCTGTGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGA	2257
Db	2426	CCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCATTTTACCACATCTGGA	2485
Qy	2258	ATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTA	2317
Db	2486	GTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAGACTCAGCCTAAACACAG	2545
Qy	2318	GTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCA	2377
Db	2546	TGGAACTGGTTTGCAAACCTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGC	2605
Qy	2378	TCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAG	2437
Db	2606	TCAACTGTACTGTGTGCTAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTGCTA	2665
Qy	2438	CGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTCGGCAGA	2497
Db	2666	GTACCATCACCACTGTCACCGGACCAAGTGCTTTTTCAGCATTCCTCTCCCTATCCGGCAGA	2725
Qy	2498	AGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCC	2557
Db	2726	AGCTATGCAGCAGCCTGGATGCCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCCC	2785
Qy	2558	AGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCA	2617
Db	2786	ATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCACTGGCG	2845
Qy	2618	TGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG	2677
Db	2846	TAATCCTGGATCTTTGGGAAGCACAGAAGTCCAGATGGAAACCTGAGCATGCTGGCAG	2905
Qy	2678	CAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGT	2737

Db 2906 CCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTACTTGGCAGCAGAAGGACAGT 2965
Qy 2738 GCTGA 2742
|||
Db 2966 ATTGA 2970

RESULT 4

CR598115

LOCUS CR598115 1852 bp mRNA linear HTC 21-JUL-2004

DEFINITION full-length cDNA clone CS0DA006YG16 of Neuroblastoma of Homo sapiens (human).

ACCESSION CR598115

VERSION CR598115.1 GI:50478922

KEYWORDS HTC; CNSLT_cDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1852)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
Faraday Avenue

REFERENCE 2 (bases 1 to 1852)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES Location/Qualifiers

source 1. .1852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA006YG16"
/tissue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 32.1%; Score 884; DB 3; Length 1852;

Best Local Similarity 100.0%; Pred. No. 7.5e-176;

Matches 884; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1869 GGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 1928
|||||

Db 1 GGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 60

Qy 1929 CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 1988
|||||

Db	61	CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG	120
Qy	1989	CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGT	2048
Db	121	CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGT	180
Qy	2049	CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT	2108
Db	181	CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT	240
Qy	2109	GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT	2168
Db	241	GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT	300
Qy	2169	GCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTGAGCTA	2228
Db	301	GCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTGAGCTA	360
Qy	2229	CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT	2288
Db	361	CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT	420
Qy	2289	CACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA	2348
Db	421	CACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA	480
Qy	2349	GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTT	2408
Db	481	GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTT	540
Qy	2409	TGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTCCCAGCCCTGGTGGGCCCCAGTGC	2468
Db	541	TGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTCCCAGCCCTGGTGGGCCCCAGTGC	600
Qy	2469	CTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAG	2528
Db	601	CTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAG	660
Qy	2529	GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCATCTCAGCTT	2588
Db	661	GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCATCTCAGCTT	720
Qy	2589	CTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTT	2648
Db	721	CTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTT	780
Qy	2649	CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGC	2708
Db	781	CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGC	840
Qy	2709	TGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	841	TGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	884

LOCUS AK018177 3866 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male medulla oblongata cDNA, RIKEN full-length enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus], full insert sequence.
 ACCESSION AK018177
 VERSION AK018177.1 GI:12857775
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3866)
 AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,

Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cdNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cdNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cdNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cdNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cdNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cdNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES

source Location/Qualifiers

1. .3866

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:6330415E02"

/db_xref="taxon:10090"

/clone="6330415E02"

/sex="male"

/tissue_type="medulla oblongata"

/clone_lib="RIKEN full-length enriched mouse cdNA library"

/dev_stage="adult"

CDS

417. .3254

/note="unnamed protein product; TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus] (SPTR|O08722, evidence: FASTY, 96.5%ID, 100%length, match=2835) putative"

/codon_start=1

/protein_id="BAB31108.1"

/db_xref="GI:12857776"

/translation="MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQVLPDSYPSAPAEQLPYFLLPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATGLRVREVQIEVSRQQVEELFGLDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLTIDHNLIIIRQARLSDTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAEWSPCSNRCGRGWQKTRTRCTNPAPLNGGAFCEGQAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLM

AEGVIVYRRNCRDFDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPDLTASAG
 IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPP
 GTYPGDFSRDTHFLHLRSASLGSQHLLGLPRDPSSSVSGTFCGLGGRLSLPGTGVSL
 VPNGAIPQKGKFDLYLHINKAESTLPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVP
 CAEVIAGDWIFQLKTQAHQGHWEVVTLDDEETLNTPCYCQLEAKSCHILLDQLGSYVF
 MGESYSRSYAVKRLQLAIFAPALCTSLEYSRLVYCLEDTPVALKEVLELERTLGGYLVE
 EPKPLLFKDSYHNLRLSLHDI PHAHWRSKLLAKYQEIPFYHVWNGSQRALHCTFTLER
 HSLASTEFTCKVCVRQVEGEGQIFQLHTTLAETPAGSLDALCSAPGNAITTLGQPYAF
 KIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ
 QDDGDLNSLASALEEMGKSEMLVAMATDGC"

ORIGIN

Query Match 31.7%; Score 872.4; DB 3; Length 3866;
 Best Local Similarity 60.6%; Pred. No. 2.3e-173;
 Matches 1646; Conservative 0; Mismatches 941; Indels 129; Gaps 8;

Qy	157	GCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAG	216
Db	540	CCCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCCTACATCGTAAAG	599
Qy	217	AACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGC	276
Db	600	AACAAGCCAGTGGAAGTGCAGTGCAGAGCCTTCCTGCCACGCAGATCTACTTCAAGTGT	659
Qy	277	AACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGT	336
Db	660	AATGGCGAGTGGGTGAGCCAGAATGACCACGTACACAGGAGAGCCTGGATGAGGCCACA	719
Qy	337	GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTTC	396
Db	720	GGCTTGCGGGTGCGAGAGGTGCAGATCGAGGTGTACGGCAGCAAGTGGAGGAAGTCTTC	779
Qy	397	GGGCTGGAGGAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGT	456
Db	780	GGGCTCGAGGACTACTGGTGCCAGTGCCTGGCCTGGAGCTCTTCGGGAAGTACCAAGAGT	839
Qy	457	CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC	516
Db	840	CGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCC	899
Qy	517	AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT	576
Db	900	AAGGAGGTACCCTTGATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAGTGCCT	959
Qy	577	CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT	636
Db	960	GTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATTGACCCGCTCAGGACACTAAC	1019
Qy	637	GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC	696
Db	1020	TTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCC	1079
Qy	697	AACTACACCTGCGTGGCCAAGAATATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTC	756
Db	1080	AACTACACCTGTGTGGCCAAGAATATCGTGGCCAAGCGCGGAGCACCGCGGCCACAGTC	1139
Qy	757	ATCGTCTACGTGAACGGTGGGTGGTGCAGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC	816

Db	1140		ATCGTCTATGTGAATGGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGC	1199
Qy	817		TGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG	876
Db	1200		TGTGGCCGAGGCTGGCAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGA	1259
Qy	877		GGCGCTTTCTGTGAGGGGAGCAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA	936
Db	1260		GGCGCCTTCTGTGAGGGACAGGCCTTCCAGAAGACAGCTTGACCACCGTGTGCCCAGTG	1319
Qy	937		GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG	996
Db	1320		GATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGG	1379
Qy	997		CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACT	1056
Db	1380		CGCAGCCGCGAGTGCATGGCACCGCCACCCAGAACGGAGGCCGTGACTGCAGCGGGACG	1439
Qy	1057		GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCC-----	1111
Db	1440		CTACTTGACTCCAAGAACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTAAAC	1499
Qy	1112		-----CTGAGGACGTGGCCCTCTATGTGGGCCTCATC	1143
Db	1500			1559
Qy	1144		GCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC---CTCATCCTCGTTTATTGCCGG	1200
Db	1560		GTGGCCGTCTTTGTGGTGGTAGCGGTTCTCATGGCCGAGGGAGTGATCGTATACGGGAGA	1619
Qy	1201		AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGT---CCATTCTCACCTCAGGCTTC	1257
Db	1620		AACTGCCGGGACTTCGACACGGACATCACCGACTCCTCTGCGGGCCCTCACTGGTGGCTTC	1679
Qy	1258		CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCATCTGCT-----CACCATC	1311
Db	1680		CACCCTGTCAACTTCAAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCC	1739
Qy	1312		CAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA-	1370
Db	1740		CCTCCAGACCTAACGGCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGAC	1799
Qy	1371		-----TGGGCCCAGCCCCAAG	1386
Db	1800			1859
Qy	1387		TTCCAGCTCACCAATGGGCACCTGCTCAG-----C	1416
Db	1860		ATCAAGGTCTATAACTCCAGCACCATCGGTTCTGGGTCTGGCCTGGCTGATGGAGCCGAC	1919
Qy	1417		CCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAG	1476
Db	1920		CTGCTGGGTGTCTCCCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTC	1979
Qy	1477		TTCGTCTCCCGCCTCTCCACCC-----AGAACTACTTCGCTCCCTGCCCCGAGGCACC	1530

Db 1980 CTGCACCTGCGCAGTGCCAGCCTTGTTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCC 2039
 Qy 1531 AGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACA 1590
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2040 AGCAGCAGTGTGAGCGGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGGCACA 2099
 Qy 1591 GGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTAC 1650
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2100 GGGGTGAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTAT 2159
 Qy 1651 CTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCCTAGCTGGCTGTCAGACCCTGCTG 1710
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2160 CTACATATCAACAAGGCCGAAAGCACCCCTCCCACTTTTCAAGGTTCCCAGACAGTATTG 2219
 Qy 1711 AGTCCCATCGTTAGCTGTGGACCCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCT 1770
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2220 AGCCCCCTCGGTGACCTGTGGGCCCCACAGGCCTACTCCTGTGCCGCCCTGTGCTCCTCACC 2279
 Qy 1771 ATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCG 1830
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2280 GTGCCCCACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCC 2339
 Qy 1831 TGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTAC 1890
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2340 CATCAGGGCCACTGGGAGGAGGTGGTGACCTTGATGAGGAGACCCTCAACACACCCTGC 2399
 Qy 1891 TACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCC 1950
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2400 TACTGCCAGCTGGAGGCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTTCTTACGTA 2459
 Qy 1951 CTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCG 2010
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2460 TTCATGGGCGAGTCTACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCC 2519
 Qy 2011 CCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCAC 2070
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2520 CCAGCCCTCTGCACCTCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCT 2579
 Qy 2071 GATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAG 2130
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2580 GTAGCACTGAAGGAGGTCTGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAG 2639
 Qy 2131 CCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTG 2190
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2640 CCCAAGCCTTTGCTCTTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCATGACATC 2699
 Qy 2191 CCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTGAGCTACCAGGAGATCCCCCTTTTATCAC 2250
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2700 CCCCATGCCCCACTGGAGGAGCAAATACTGGCCAAGTACCAGGAGATTCCCTTCTACCAC 2759
 Qy 2251 ATCTGGAATGGCAGCGAGCGGTACTTGCAGTGCACCTTACCCTGGAGCGTGTGAGCCCC 2310
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2760 GTCTGGAATGGCAGCCAGAGAGCCCTGCACTGCACTTTTACCCTGGAGAGGCATAGCCTG 2819
 Qy 2311 AGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGC 2370
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2820 GCCTCCACGGAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTGGAAGGGGAAGGCCAGATT 2879

Qy 2371 TTCAGCATCAACTTCAACATCACC---AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTG 2427
 ||| | || | || | || | || | ||
 Db 2880 TTCCAGCTGCACACAACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCT 2939

Qy 2428 GAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCCTTCCTC 2487
 | | | || | ||| ||| ||| ||| ||| ||| ||| ||| ||| |
 Db 2940 GCCCCGGGCAATGCCATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCC 2999

Qy 2488 ATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGG 2547
 || || || |||| |||| |||| || || || || || || || ||
 Db 3000 ATCCGCCAAAAGATCTGCAGCAGCCTGGACGCCCCCGACTCCCGGGGCAACGACTGGAGG 3059

Qy 2548 ACTCTGGCCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGC 2607
 |||| |||| || |||| || || || || || || || || || ||
 Db 3060 CTGTTGGCGCAGAAGCTGTCCATGGACCGGTACCTAACTACTTCGCCACCAAAGCTAGT 3119

Qy 2608 CCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGC 2667
 |||| || | || | ||| |||| || |||| || || || || ||
 Db 3120 CCCACAGGTGTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAAC 3179

Qy 2668 CAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTG 2727
 |||| || || | |||| || || || || || || || || ||
 Db 3180 AGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACA 3239

Qy 2728 GAGGCTGAGTGCTGAG 2743
 || | || |||||
 Db 3240 GATGGCGATTGCTGAG 3255

RESULT 6

AY406492

LOCUS AY406492 2802 bp DNA linear GSS 15-DEC-2003

DEFINITION Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY406492

VERSION AY406492.1 GI:39762466

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 2802)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2802)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers
source 1. .2802
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1. .>2802
/gene="UNC5C"
/locus_tag="HCM2575"

ORIGIN

Query Match 29.6%; Score 814; DB 9; Length 2802;
Best Local Similarity 53.8%; Pred. No. 4.7e-161;
Matches 1477; Conservative 0; Mismatches 1200; Indels 69; Gaps 6;

```
Qy      57  GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 116
      || || || | | | ||| || | | | | | | | | | |
Db      66  GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125

Qy     117  TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 176
      || | | | | | | | | | | | | | | | | | |
Db     126  TGATGACTTTTTTCATGAACCTCCAGAACTTTTCCTTCGGATCCACCTGAGCCTCTGCC 185

Qy     177  CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 236
      || |||| | || || |||| | || || || |||| |||| ||
Db     186  ACATTTTCCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245

Qy     237  GTGCAAGGCC-----GTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGT 290
      || || || | |
Db     246  CTGTAAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 305

Qy     291  GCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCAT 350
Db     306  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 365

Qy     351  GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGCGGCTGGAGGAATA 410
Db     366  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 425

Qy     411  CTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACAT 470
Db     426  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 485

Qy     471  CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT 530
      || |||| || |||| || || || || || || || || || ||
Db     486  NNNNNNNNNNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTT 545

Qy     531  GGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA 590
      ||| |||| || | || | || || |||| || || |||| || ||
Db     546  GGAACAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNN 605

Qy     591  GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG 650
      || || ||| | | || || || || || || || || || ||
Db     606  NNNNNNGAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT 665
```

Qy	651	GGAGCACAGCCTGGTGGTGGCGACAGGCCCGCCTTGCTGACACGGCCAACCTACACCTGCGT	710
Db	666	TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTATACCTGTGT	725
Qy	711	GGCCAAGAACATCGTGGGCACGTCGCCGACGCGCTCCGCTGCTGTCATCGTCTACGTGAA	770
Db	726	TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACCTGCCACTGTCATAGTCTATGTCAA	785
Qy	771	CGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG	830
Db	786	CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTA	845
Qy	831	GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA	890
Db	846	TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCCTCAATGGGGGTGCCTTCTGTGA	905
Qy	891	GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAG	950
Db	906	AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGAC	965
Qy	951	CCCGTGGAGCAAGTGGTGGCGCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTG	1010
Db	966	GCCATGGAGCAAGTGGTCTACTTGTGGAAGTGAAGTGCACCCACTGGCGCAGGAGGGAGTG	1025
Qy	1011	CTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCG	1070
Db	1026	CACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGATGGCCTCGTCTTGCAATCCAA	1085
Qy	1071	CAACTGTACCACTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTA	1130
Db	1086	GAAGTGAAGTGAAGTGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTA	1145
Qy	1131	TGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTGCTTGTCTCTCATCCT	1187
Db	1146	TGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTT	1205
Qy	1188	CGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC	1247
Db	1206	TGTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAA	1265
Qy	1248	CTCAGGCTTCCAGCCCGTCAGCATCAAGCCAGCAAAGCAGACAACCCCCATCTGCTCAC	1307
Db	1266	TGGGGGCTTTTCAGCCTGTGAACATCAAG-----GCAGCAAGACAAGATCTGCTGGC	1316
Qy	1308	CATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCA	1367
Db	1317	TGTACCCCCAGACCTCACGTGAGTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCA	1376
Qy	1368	GGATG-----GGCCAGCCCCAAGTTCCAGCTCACCAAT-----GGGCACCTGCTCAGCCC	1418
Db	1377	TGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCAACCT	1436
Qy	1419	CCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTT	1478
Db	1437	GAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGAGTT	1496

Qy	1479	CGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCGCAACAT	1538
Db	1497	TACGTCCAAGCTGTCCCCTCAGATGACCCAGTCATTGTTGGAGAATGAAGCCCTCAGCCT	1556
Qy	1539	GA-----CCTATGGGACCTTCAA	1556
Db	1557	GAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCNN	1616
Qy	1557	CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA	1616
Db	1617	NTNGCTGGGNNNNCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCTGG	1676
Qy	1617	TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT	1676
Db	1677	GGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAACTAT	1736
Qy	1677	GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC	1736
Db	1737	GAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCC	1796
Qy	1737	TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC	1796
Db	1797	AGGAGCTCTGCTCACCCGCCCGCTCGTCCTTACTATGCATCACTGCGCAGACCCCAATAC	1856
Qy	1797	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	1856
Db	1857	CGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAGGATGTGGT	1916
Qy	1857	GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	1916
Db	1917	GGTGGTCGGGGAGGAAAACCTCACCACCCCCTGCTACATTGAGCTGGATGCAGAGGCCTG	1976
Qy	1917	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	1976
Db	1977	CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACCAAAGC	2036
Qy	1977	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA	2036
Db	2037	GGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCCTGTGCTGCTCCTCGCTGGAGTA	2096
Qy	2037	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	2096
Db	2097	CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT	2156
Qy	2097	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTTGCACCTTCAAGGACAG	2156
Db	2157	TGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAG	2216
Qy	2157	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT	2216
Db	2217	CACCCACAACCTGCGCCTGTCAATTACGATATCGCCATTCCCTCTGGAAGAGCAAATT	2276
Qy	2217	CCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT	2276
Db	2277	GCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAACCT	2336
Qy	2277	GCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT	2336

Db 2337 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396
 Qy 2337 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2396
 Db 2397 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGA 2456
 Qy 2397 GGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2456
 Db 2457 GGAACCTACTGGCATCGATTGCGGCTGCTGGATCCTGCGAACACCATCACCGGTGAC 2516
 Qy 2457 GGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTGCGCAGAAGATAATTTCCAGCCTGGA 2516
 Db 2517 GGGGCCCAGTGCTTTTCCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576
 Qy 2517 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAG 2576
 Db 2577 TGCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCNNNNNNNNNNNNNN 2636
 Qy 2577 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2636
 Db 2637 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696
 Qy 2637 GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2696
 Db 2697 AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756
 Qy 2697 CCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGA 2742
 Db 2757 AAGACATGAAACGGTGGTGTCTTAGCAGCAGAAGGGCAGTATTAA 2802

RESULT 7

AY411747

LOCUS AY411747 2532 bp DNA linear GSS 12-DEC-2003

DEFINITION Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY411747

VERSION AY411747.1 GI:39767715

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2532)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2532)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers
source 1. .2532
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>2532
/locus_tag="HCM4327"

ORIGIN

Query Match 29.4%; Score 810.4; DB 9; Length 2532;
Best Local Similarity 61.1%; Pred. No. 2.7e-160;
Matches 1540; Conservative 0; Mismatches 851; Indels 129; Gaps 9;

```
Qy      352 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATAC 411
      ||||| |  ||  | || ||  ||||| || || || || ||||| |||
Db      13  GAGGTGCAGATCGAGGTGTGCGGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72

Qy      412 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 471
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      73  TGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC 132

Qy      472 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 531
      ||||| |||  ||||| ||||| ||||| ||||| ||||| |||||
Db      133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCTG 192

Qy      532 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 591
      || ||| |  | ||| |||| ||||| ||||| ||||| ||||| |||||
Db      193 GACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAA 252

Qy      592 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 651
      ||||| ||| ||||| |  ||||| ||| |||| |||| |  ||||
Db      253 TGGCTCAAGAATGAGGATGTCATCGACCCACCCAGGACACCAACTTCCTGCTCACCATC 312

Qy      652 GAGCACAGCCTGGTGGTGGCAGAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 711
      || |||| ||| |  || ||||| ||||| ||||| ||||| ||||| |||||
Db      313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTGCGGACACTGCCAACTATACCTGCGTG 372

Qy      712 GCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCTACGTGAAC 771
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCCTGCCACCGTCATCGTCTACGTGAAT 432

Qy      772 GGTGGGTGGTGCAGCTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 831
      || || |||| |  ||| ||||| ||||| ||||| ||||| ||||| |||||
Db      433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTACCCTGCTCCAACCGCTGTGGCCGAGGCTGG 492

Qy      832 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 891
      ||||| || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      493 CAGAAGCGCACCCGGACCTGCACCAACCCGCTCCACTCAACGGAGGGGCCTTCTGCGAG 552

Qy      892 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC 951
      || |||  ||||| || ||||| ||||| ||||| ||||| ||||| |||||
```


Db	553	GGCCAGGCATTCCAGAAGACCGCCTGCACCACCATCTGCCCAGTCGATGGGGCGTGGACG	612
Qy	952	CCGTGGAGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC	1011
Db	613	GAGTGGAGCAAGTGGTTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGC	672
Qy	1012	TCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC	1071
Db	673	ATGGCGCCCCACCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAG	732
Qy	1072	AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCT-----	1107
Db	733	AACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACCTCTAAGCGACCCCAACAGCCAC	792
Qy	1108	-----GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC---ATCGCCGTGGCCGTC	1155
Db	793	CTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTG	852
Qy	1156	TGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG	1215
Db	853	GTCGTGGCAATCCTCATGGCGGTGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTC	912
Qy	1216	GACTCAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATC	1272
Db	913	GACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTT	972
Qy	1273	AAGCCCAGCAAAGCAGACAACCCCATCTGCT-----CACCATCCAGCCGGACCTCAGC	1326
Db	973	AAGACGGCAAGGCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA	1032
Qy	1327	ACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG-----	1374
Db	1033	GCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATC	1092
Qy	1375	-----CCCAGCCCCAAGTTCCAGCTCACCA--	1399
Db	1093	CCCATGACCAACTCTCCTCTGCTGGACCCCTTACCAGCCTTAAGGTCAAGGTCTACAGC	1152
Qy	1400	-----ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCC-----ACACA	1440
Db	1153	TCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTG	1212
Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAG	1500
Db	1213	CCGCCTGGCACATAACCTAGCGATTTCGCCCCGGGACACCCACTTCTGCACCTGCGCAGC	1272
Qy	1501	AACTACTTCCGCTC-----CCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1273	GCCAGCCTCGGTTCCACGAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGC	1332
Qy	1546	GGGACCTTCAACTTCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1333	GGCACCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTGAGCTTGCTG	1392
Qy	1606	ATCCCCCAGATGCCATAACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1393	GTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAG	1452

Qy	1666	CCGGAAGACGTGAGGTTGCCCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1725
Db	1453	GCAGAAAGTACCCTCCCCTTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACC	1512
Qy	1726	TGTGGACCCCTGGCGTCTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1513	TGTGGACCCACAGGCCCTCTGCTGTGCCGCCCGTCTATCCTCACCATGCCCCACTGTGCC	1572
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGGAGCTGG	1845
Db	1573	GAAGTCAGTGCCCGTGACTGGATCTTTTCAGCTCAAGACCCAGGCCACCAGGGCCACTGG	1632
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1633	GAGGAGGTGGTGACCTGGATGAGGAGACCCCTGAACACACCCTGCTACTGCCAGCTGGAG	1692
Qy	1906	GCCAGTGCCCTGCTACGTCTTCACCGAGCAGCTGGGGCCGCTTTGCCCTGGTGAGAGGCC	1965
Db	1693	CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTACGGGCGAGTCC	1752
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1753	TATTCCCCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC	1812
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1813	TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG	1872
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2145
Db	1873	GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG	1932
Qy	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205
Db	1933	TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCATGCCCATTTGG	1992
Qy	2206	AAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCCTTTTATCACATCTGGAATGGCAG	2265
Db	1993	AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCCTTCTATCACATTTGGAGTGGCAGC	2052
Qy	2266	CAGCGGTACTTGCACCTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG	2325
Db	2053	CAGAAGGCCCTCCACTGCACCTTACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC	2112
Qy	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
Db	2113	ACCTGCAAGATCTGCGTGCGGCAAGTGAAGGGGAGGGCCAGATATTCCAGCTGCATACC	2172
Qy	2386	AACATCACCAAG---GACACAAGGTTTGTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG	2442
Db	2173	ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACT	2232
Qy	2443	GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCCTTCCCTATTTCGGCAGAAGATA	2502
Db	2233	GTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATA	2292

Qy 2503 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2562
 ||||| .|| | ||| ||||| ||||| |||||
 Db 2293 TGCAACAGCCTAGATGCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG 2352

Qy 2563 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2622
 ||| ||||| | | || | | ||||| |||| | ||||| | |||||
 Db 2353 CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAGCGAGCCCCACGGGTGTGATC 2412

Qy 2623 CTCAACCTGTGGGAGGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2682
 || |||| ||||| || | ||| | | ||||| | ||||| ||
 Db 2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC 2472

Qy 2683 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGA 2742
 ||| | ||||| || || | | || | || | |||||
 Db 2473 TTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2532

RESULT 8

AY411749

LOCUS AY411749 2532 bp DNA linear GSS 12-DEC-2003

DEFINITION Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY411749

VERSION AY411749.1 GI:39767717

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2532)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2532)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1. .2532
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1. .>2532
 /locus_tag="HCM4327"

ORIGIN

Query Match 28.4%; Score 780.4; DB 9; Length 2532;
Best Local Similarity 60.1%; Pred. No. 5.8e-154;
Matches 1515; Conservative 0; Mismatches 876; Indels 129; Gaps 8;

```
Qy      352 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCTGGGCTGGAGGAATAC 411
        ||||| |  ||  | || ||| ||||| || ||| |  | ||||| ||||| |||
Db      13  GAGGTGCAGATCGAGGTGTACGGCAGCAAGTGGAGGAACCTCTTCGGGCTCGAGGACTAC 72

Qy      412 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 471
        ||||| ||||| ||||| ||||| ||||| || ||||| ||||| ||||| |||||
Db      73  TGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACACCAAGAGTCGCCGAGCCTACATC 132

Qy      472 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 531
        ||||| |||  ||||| ||||| || ||||| ||||| ||||| ||||| |||||
Db      133 CGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCCAAGGAGGTACCCTTG 192

Qy      532 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 591
        || ||| |  || | |||| ||||| ||||| ||||| || |||  || ||||| |||
Db      193 GATCATGAGGTCCTTCTGCAGTGCCGCCACCGGAGGGAGTGCCTGTGGCTGAGGTGGAA 252

Qy      592 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 651
        ||||| ||| || || |  | ||||| || ||||| || || |  |||||
Db      253 TGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAACTTCCTGCTCACCATT 312

Qy      652 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACCTACACCTGCGTG 711
        || |||| |||  |  | || |||| ||||| || ||||| ||||| ||||| |||
Db      313 GACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCCAACCTACACCTGTGTG 372

Qy      712 GCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTATCGTCTACGTGAAC 771
        ||||| ||||| ||||| ||||| || || || || || ||||| ||||| |||||
Db      373 GCCAAGAATATCGTGGCCAAGCGCCGAGCACCACGGCCACAGTCATCGTCTATGTGAAT 432

Qy      772 GGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 831
        || || ||||| |  || | ||||| ||||| ||||| || ||||| || |||||
Db      433 GGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGCTGTGGCCGAGGCTGG 492

Qy      832 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 891
        ||||| || |  |||| ||||| ||||| || || || ||||| || ||||| |||||
Db      493 CAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGGCGCCTTCTGTGAG 552

Qy      892 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGC 951
        || |||  ||||| ||||| || ||||| ||||| || || ||||| |||||
Db      553 GGACAGGCCTTCCAGAAGACAGCTTGACCACCCTGTGCCAGTGGATGGAGCGTGGACC 612

Qy      952 CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC 1011
        ||||| ||||| ||||| ||||| || || || ||||| ||||| |||||
Db      613 GAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAGTGC 672

Qy      1012 TCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC 1071
        |  || ||||| ||||| ||||| || || || || || |||||
Db      673 ATGGCACCGCCACCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCCAAG 732

Qy      1072 AACTGTACCACTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC----- 1119
        ||||| ||  || || || || |  |  |  |||||
Db      733 AACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGCCAC 792

Qy      1120 -----GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGC 1158
```

Db	793	CCCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTTGTGGTGGCCGTCTTTGTG	852
Qy	1159	CTGGTCCTGCTGCTGCTTGTCTCT---CATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTG	1215
Db	853	GTGGTAGCGGTTCTCATGGCCGTGGGAGTGATCGTATACCGGAGAACTGCCGGGACTTC	912
Qy	1216	GACTCAGATGTGGCTGACTCGT---CCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATC	1272
Db	913	GACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGCTTCCACCCTGTCAACTTC	972
Qy	1273	AAGCCCAGCAAAGCAGACAACCCCATCTGCT-----CACCATCCAGCCGGACCTCAGC	1326
Db	973	AAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCCCCTCCAGACCTAACG	1032
Qy	1327	ACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGG-----	1373
Db	1033	GCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGACTCCGCCGACAAGATC	1092
Qy	1374	-----GCCCAGCCCCAAGTTCAGCTCACCAAT	1401
Db	1093	CCCATGACTAATTGCCCCCTGCTGGATCCCCTGCCCAGCCTCAAGATCAAGGTCTATAAC	1152
Qy	1402	GGGCACCTGCTCAG-----CCCCCTGGGTGGCGGC	1431
Db	1153	TCCAGCACCATCGGTTCTGGGTCTGGCCTGGCTGATGGAGCCGACCTGCTGGGTGTCCTC	1212
Qy	1432	CGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTC	1491
Db	1213	CCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTCTGCACCTGCGCAGT	1272
Qy	1492	TCCACCCAGAACT-----ACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1273	GCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCCAGCAGCAGTGTGAGC	1332
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1333	GGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGCACAGGGGTGAGCCTGTTG	1392
Qy	1606	ATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1393	GTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTATCTACATATCAACAAG	1452
Qy	1666	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1453	GCCGAAAGCACCCCTCCCACCTTTCAGAAGGTTCCAGACAGTATTGAGCCCCTCGGTGACC	1512
Qy	1726	TGTGGACCCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1513	TGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACCGTGCCCCACTGTGCT	1572
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1573	GAAGTCATCGCTGGAGACTGGATCTTTAGCTCAAGACCCAGGCCCATCAGGGCCACTGG	1632
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1905

Db	1633	GAGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACACCCTGCTACTGCCAGCTGGAG	1692
Qy	1906	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCGCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1693	GCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTACCTACGTATTCATGGGCGAGTCC	1752
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1753	TACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCCCCAGCCCTCTGCACC	1812
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1813	TCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCTGTAGCACTGAAGGAG	1872
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2145
Db	1873	GTCTTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAGCCCAAGCCTTTGCTC	1932
Qy	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205
Db	1933	TTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCATGACATCCCCATGCCCACTGG	1992
Qy	2206	AAGAGTAAGCTCCTTGTCTAGCTACCAGGAGATCCCCCTTTTATCACATCTGGAATGGCAG	2265
Db	1993	AGGAGCAAACCTACTGGCCAAGTACCAGGAGATTCCCTTCTACCACGTCTGGAATGGCAGC	2052
Qy	2266	CAGCGGTACTTGCACCTGCACCTTACCCCTGGAGCGTGTCTAGCCCCAGCACTAGTGACCTG	2325
Db	2053	CAGAGAGCCCTGCACCTGCACCTTACCCCTGGAGAGGCATAGCCTGGCCTCCACGGAGTTC	2112
Qy	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
Db	2113	ACCTGTAAGGTCTGCGTGCAGGTCGAAGGGGAAGGCCAGATTTTCCAGCTGCACACA	2172
Qy	2386	AACATCACC---AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG	2442
Db	2173	ACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCTGCCCCGGGCAATGCC	2232
Qy	2443	GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTGGCGAGAAGATA	2502
Db	2233	ATCACCAACCAGCTGGGACCCCTATGCCTTCAAGATACCCCTGTCCATCCGCCAAAAGATC	2292
Qy	2503	ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA	2562
Db	2293	TGCAGCAGCCTGGACGCCCCCAACTCCCGGGGCAACGACTGGAGGCTGTTGGCGCAGAAG	2352
Qy	2563	CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCCTCCAAGCCCAGCCCCACAGCCATGATC	2622
Db	2353	CTGTCCATGGACCGGTACCTAACTACTTCGCCACCAAAGCTAGTCCCACAGGTGTCATC	2412
Qy	2623	CTCAACCTGTGGGAGGCGCGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA	2682
Db	2413	TTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAACAGCCTGGCCAGTGCC	2472
Qy	2683	GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGA	2742
Db	2473	TTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACAGATGGCGATTGCTGA	2532

RESULT 9

BI758231

LOCUS BI758231 1034 bp mRNA linear EST 25-SEP-2001

DEFINITION 603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5', mRNA sequence.

ACCESSION BI758231

VERSION BI758231.1 GI:15749809

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1034)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11501 row: g column: 20

High quality sequence stop: 793.

FEATURES

source

Location/Qualifiers

1. .1034

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5200171"

/lab_host="DH10B"

/clone_lib="NIH_MGC_114"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 27.9%; Score 768.6; DB 4; Length 1034;

Best Local Similarity 91.2%; Pred. No. 1.6e-151;

Matches 918; Conservative 0; Mismatches 74; Indels 15; Gaps 9;

Qy 150 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 209

|||||

Db 1 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 60

Qy 210 CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTT 269

Db	61		CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT	120
Qy	270		CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG	329
Db	121		CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG	180
Qy	330		GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAA	389
Db	181		GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAA	240
Qy	390		GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC	449
Db	241		GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC	300
Qy	450		CAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAAGCTTCGAGCAGGAGCC	509
Db	301		CAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAAGCTTCGAGCAGGAGCC	360
Qy	510		GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG	569
Db	361		GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG	420
Qy	570		CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA	629
Db	421		CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA	480
Qy	630		CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA	689
Db	481		CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA	540
Qy	690		CACGGCCAACCTACACCTGCGTGGCCAAGAACATCGTGGCACGTGCGCCGAGCGCCTCCGC	749
Db	541		CACGGACAACCTACACCTGCGTGGCCAAGAACATCGTGGCACGTGCGCCGAGCGCCTCCGC	600
Qy	750		TGCTGTCATCGTCTACGTGAACGGTGGGTGGTTCGA-CGTGGACCGAGTGGTCCGTCTGCA	808
Db	601		TGCTGTCATCGTCTACGTGAACGGTGGGTGGTTCGACCGTGGACCGAGTGGTCCGTCTGC-	659
Qy	809		GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	868
Db	660		GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACAACCCGGTGCCTC	719
Qy	869		TCAACGGGGGCGCTTTCTGTGA-GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTG	927
Db	720		TCAACGGGGGCGCTTTCTGTGAGGGGGCAGAATGTCCAGAAAGCAGC---TGCGCCACCT	776
Qy	928		TGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTGGCCTGT---GGGCTGGACT	985
Db	777		GTGCCAGTGGACGGCAGCTGTAGCACGTGGAGCCAGTGGTGGCCTGTTGGGCTTGGCTT	836
Qy	986		GCACCCACT--GGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGA	1043
Db	837		GCACCCACTTGGGCGGAGCCGGAGTGCTCTGAACCCAGCACCCCGGCACGGCAGGGGGAG	896
Qy	1044		GTG---CCAGGGCACTGACCTGGACACCCGCAACTGTACCAAGTACCTCTGTGTACACAG	1100

Db 897 GTGTGCCCAGGTCACCTGGACCTGGCACCCGGGA-TGGTCCAGTGAGCTCTGTGT-CCCAC 954

Qy 1101 TGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCG 1147
 | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 955 GGGTTCTGGCCCTGAGGACTTGGCCTCCTATGTGGGCCTCATCCCG 1001

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nhgri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the T.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 68 Row: i Column: 2

This clone was selected for full length sequencing because it
passed the following selection criteria: GenomeScan gene prediction
This clone has the following problem: frame shifted.

FEATURES	Location/Qualifiers
source	1. .1532
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:5166762"
	/tissue_type="Brain, adult medulla"
	/clone_lib="NIH_MGC_119"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 26.8%; Score 736.6; DB 3; Length 1532;
Best Local Similarity 84.2%; Pred. No. 9.7e-145;
Matches 917; Conservative 0; Mismatches 4; Indels 168; Gaps 1;

Qy	337	GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTC	396
Db	612	GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTC	671
Qy	397	GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT	456
Db	672	GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT	731
Qy	457	CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC	516
Db	732	CAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC	791
Qy	517	AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT	576
Db	792	AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT	851
Qy	577	CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAAT	636
Db	852	CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAAT	911
Qy	637	GTATACATCACGCGGGAGCACAGCCTGGTGGTGGCAGAGCCCGCCTTGCTGACACGGCC	696
Db	912	GTATACATCACGCGGGAGCACAGCCTGGTGGTGGCAGAGCCCGCCTTGCTGACACGGCC	971

Qy	697	AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC	756
Db	972	AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC	1031
Qy	757	ATCGTCTACGTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC	816
Db	1032	ATCGTCTACGTG-----	1043
Qy	817	TGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG	876
Db	1044	-----	1043
Qy	877	GGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA	936
Db	1044	-----	1043
Qy	937	GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG	996
Db	1044	GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG	1103
Qy	997	CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGGCACT	1056
Db	1104	CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGGCACT	1163
Qy	1057	GACCTGGACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAG	1116
Db	1164	GACCTGGACACCCGCAACTGTACCAAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAG	1223
Qy	1117	GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTGCTT	1176
Db	1224	GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTGCTT	1283
Qy	1177	GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG	1236
Db	1284	GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG	1343
Qy	1237	TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC	1296
Db	1344	TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC	1403
Qy	1297	CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTC	1356
Db	1404	CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTC	1463
Qy	1357	TGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC	1416
Db	1464	TGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC	1523
Qy	1417	CCCCTGGGT	1425
Db	1524	CCCCTGGGT	1532

RESULT 11

AI951556

LOCUS

AI951556

788 bp

mRNA

linear

EST 09-MAR-2000

DEFINITION wv36f04.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:2531647. 3' similar to TR:008721 008721 TRANSMEMBRANE RECEPTOR UNC5H1. ;, mRNA sequence.

ACCESSION AI951556

VERSION AI951556.1 GI:5743866

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 788)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1125 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 446.

FEATURES
source Location/Qualifiers
1. .788
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2531647"
/tissue_type="fibrotheoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI_CGAP_Ov18"
/note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGACATTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 26.7%; Score 735.2; DB 1; Length 788;
Best Local Similarity 96.5%; Pred. No. 1.8e-144;
Matches 749; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1469 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCGCTCCCTGCCCCGAGGCA 1528
|||||

Db 1 CCGAGGAGTTCGTCTCCCGCCTCTCCGCCCAGAACTACTTCGCTCCCTGCCCCGAGGCA 60

Qy 1529 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1588

```

      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 120

Qy      1589 CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1648
      |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CAGGAATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 180

Qy      1649 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 1708
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 240

Qy      1709 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG 1768
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTTCTGCTCACCCGGCCAGTCATCCTGG 300

Qy      1769 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1828
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 360

Qy      1829 CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCT 1888
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCT 420

Qy      1889 ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG 1948
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCTATG 480

Qy      1949 CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 2008
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 540

Qy      2009 CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCC 2068
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATGACACTC 600

Qy      2069 ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGG 2128
      ||||||||||||| | ||||||||||||||||||||||| |||||||||||||||
Db      601 ACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGG 660

Qy      2129 AGCCACGGGTCTCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATG 2188
      |||||| ||| ||||||||||||||||||||||| |||||||||||||||
Db      661 AGCCACTGGTACTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATNCATCCACGATG 720

Qy      2189 TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTT 2244
      ||||||||| || | ||||||||| ||||||||||| ||||||| |||
Db      721 TGCCCAGCTNCCNTGTGGAGAGTAAGCTTCTTGTGCTAGCTACCCAGAGATCCNCTAT 776

```

RESULT 12

BX348193/c

LOCUS BX348193 796 bp mRNA linear EST 08-APR-2004

DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.

ACCESSION BX348193

VERSION BX348193.2 GI:46286231

KEYWORDS EST.

SOURCE Homo sapiens (human)

[illegible]

Db 556 TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT 497
 Qy 1978 GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTAC 2037
 |||||
 Db 496 GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTAC 437
 Qy 2038 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 2097
 |||||
 Db 436 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 377
 Qy 2098 GAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGT 2157
 |||||
 Db 376 GAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGT 317
 Qy 2158 TACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTC 2217
 |||||
 Db 316 TACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTC 257
 Qy 2218 CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCAGCAGCGGTACTTG 2277
 |||||
 Db 256 CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCAGCAGCGGTACTTG 197
 Qy 2278 CACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 2337
 |||||
 Db 196 CACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 137
 Qy 2338 TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG 2397
 |||||
 Db 136 TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG 77
 Qy 2398 GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTG 2457
 |||||
 Db 76 GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGNCCCTGGT 17
 Qy 2458 GGCCCC 2463
 || |||
 Db 16 GGGCCC 11

RESULT 13

BI818609

LOCUS BI818609 818 bp mRNA linear EST 04-OCT-2001

DEFINITION 603033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5', mRNA sequence.

ACCESSION BI818609

VERSION BI818609.1 GI:15929902

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 818)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11434 row: 1 column: 16
 High quality sequence stop: 744.

FEATURES
 source Location/Qualifiers
 1. .818
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5174559"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.0%; Score 716.4; DB 4; Length 818;
 Best Local Similarity 97.7%; Pred. No: 1.7e-140;
 Matches 748; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

Qy	43	GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG	102
Db	1	GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG	60
Qy	103	CTCCGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC	162
Db	61	CTCCGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC	120
Qy	163	CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG	222
Db	121	CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG	180
Qy	223	CCAGTGCTGCTTGTGTGCAAGGCCGTGCCC GCCACGCAGATCTTCTTCAAGTGCAACGGG	282
Db	181	CCAGTGCTGCTTGTGTGCAAGGCCGTGCCC GCCACGCAGATCTTCTTCAAGTGCAACGGG	240
Qy	283	GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG	342
Db	241	GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG	300
Qy	343	CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTCTGGGCTG	402
Db	301	CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTCTGGGCTG	360

Qy 403 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 462
 |||
 Db 361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420
 Qy 463 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 522
 |||
 Db 421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480
 Qy 523 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 582
 |||
 Db 481 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 540
 Qy 583 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 642
 |||
 Db 541 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 600
 Qy 643 ATCACGCGGGAGCACAGCCTGGTGGTGGCAGGCCCCGCCTTGCTGACACGGCCAACCTAC 702
 |||
 Db 601 ATCACGCGGGAGCACAGCCTGGTGGTGGCAGGCCCCGCCTTGCTGACACGGCCAACCTAC 660
 Qy 703 ACCTGCGTGGCCAAGAACATCGTGGCACGTC-GCCGCAGCGCCTCCGCTGCTGTCATCGT 761
 |||
 Db 661 ACCTGCGTGGCCAAGAACATCGTGGCACGTCAGCCGCAGGGCCTCCGCTGCTGTCATCGT 720
 Qy 762 CTACGTGAACGGTGGGTGGTTCGACG-TGGACCGAGTGGTCCGTCTG 806
 |||
 Db 721 CTACGTGGACGGCAGCTGGAGCCCGTTGGAGCCAGTGGTTCGGGCTG 766

RESULT 14

BX364574

LOCUS BX364574 853 bp mRNA linear EST 08-APR-2004

DEFINITION BX364574 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.

ACCESSION BX364574

VERSION BX364574.2 GI:46286957

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 853)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On May 5, 2003 this sequence version replaced gi:30368812.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3529.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?s=CS1AC006ZF06QP1&c=3529.f>.

FEATURES
source Location/Qualifiers.
1. .853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC023YB10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 25.1%; Score 692; DB 5; Length 853;

Best Local Similarity 99.0%; Pred. No. 2.4e-135;

Matches 703; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```
Qy      1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60
      ||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 CCGCRGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 204

Qy      61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 264

Qy     121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
      ||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     265 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCAACCCGGACCTGCTTCCCCAC 324

Qy     181 TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 384

Qy     241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 444

Qy     301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC 360
      |||||||||||||||||:|||||||||||||||||||||||||
Db     445 GACCACGTGATCGAGCGCA-SACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC 503

Qy     361 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 420
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     504 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 563

Qy     421 TCGGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     564 TCGGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 623

Qy     481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     624 TATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 683

Qy     541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
      |||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db 684 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 743

Qy 601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 744 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 803

Qy 661 CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT 710
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 804 CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT 853

RESULT 15

AY411748

LOCUS AY411748 2532 bp DNA linear GSS 12-DEC-2003

DEFINITION Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY411748

VERSION AY411748.1 GI:39767716

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 2532)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2532)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1. .2532
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 gene <1..>2532
 /locus_tag="HCM4327"

ORIGIN

Query Match 24.6%; Score 678.2; DB 9; Length 2532;
 Best Local Similarity 54.4%; Pred. No. 2.2e-132;
 Matches 1370; Conservative 0; Mismatches 1021; Indels 129; Gaps 9;

Qy 352 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCTGGGCTGGAGGAATAC 411
 ||||| | || | || || ||||| ||| || | || ||||| |||

Db	13	GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC	72
Qy	412	TGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 	471
Db	73	TGGTGCCAGTGCCTGGCCTGGAGCTCTGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC	132
Qy	472	CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 	531
Db	133	CGCATCGCCTNNN	192
Qy	532	GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 	591
Db	193	NNNGTGGAA	252
Qy	592	TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 	651
Db	253	TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCTGCTCACCATC	312
Qy	652	GAGCACAGCCTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAATACTACACCTGCGTG 	711
Db	313	GACCACAACCTCATCATCCGCCAGGCCCGCCTGTGCGACACTGCCAATACTATACCTGCGTG	372
Qy	712	GCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 	771
Db	373	GCCAAGAACATCGTGGCAAACGCCGGAGCACCCTGCCACCGTCATCGTCTACGTGAAT	432
Qy	772	GGTGGGTGGTGCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 	831
Db	433	GGCGGCTGGTCCAGCTGGGCAGAGTGGTGCCTGCTCCAACCGCTGTGGNNGANNCTGG	492
Qy	832	CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 	891
Db	493	CAGAAGCGCACCCGGACCTGCACCAACCCGCCCACTCAACGGAGGGNNNTCTGCGAG	552
Qy	892	GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGC 	951
Db	553	GGCCANNTCGATGGGGCGTGGACG	612
Qy	952	CCGTGGAGCAAGTGGTCCGCCTGTGGGCTGGACTGCACCACTGGCGGAGCCGTGAGTGC 	1011
Db	613	GAGTGGAGCAAGTGGTCAACCTGCAGCACTGAGTGTGCCACTGGCGTAGCCGCGAGTGC	672
Qy	1012	TCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC 	1071
Db	673	ATGGCGCCCCAACCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAG	732
Qy	1072	AAGTGTACCAAGTACCTCTGTGTACACAGTGTCT----- 	1107
Db	733	AAGTGCACAGATGGGCTGTGCATGAAAATAAGAAAACTCTAAGCGACCCCAACAGCCAC	792
Qy	1108	-----GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC---ATCGCCGTGGCCGTC 	1155
Db	793	CTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCGTCTTCNTG	852
Qy	1156	TGCTTGGTCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG 	1215
Db	853	GTCNTGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTC	912

Qy	1216	GACTCAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCACGCCCGTCAGCATC	1272
Db	913	GACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTT	972
Qy	1273	AAGCCCAGCAAAGCAGACAACCCCCATCTGCT-----CACCATCCAGCCGGACCTCAGC	1326
Db	973	AAGACGGCAAGGCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA	1032
Qy	1327	ACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG-----	1374
Db	1033	GCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATC	1092
Qy	1375	-----CCCAGCCCCAAGTTCCAGCTCACCA--	1399
Db	1093	CCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGC	1152
Qy	1400	-----ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCC-----ACACA	1440
Db	1153	TCCAGCACTACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGAGTCTTG	1212
Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Db	1213	CCGCCTGGCACATAACCCTAGCGATTTGCCCCGGGACACCCACTTCCTGCACCTGCGCAGC	1272
Qy	1501	AACTACTTCCGCTC-----CCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1273	GCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGNNAGCAGCGTCAGC	1332
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1333	GGCACCTTNNGGGTCAGCTTGCTG	1392
Qy	1606	ATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1393	GTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAG	1452
Qy	1666	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTGACACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1453	GCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCCTCAGTGACC	1512
Qy	1726	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1513	TGTGGACCCACAGGCCTCCTGCTGTGCCACCCCGTCATCCTCACCATGCCCCACTGTGCC	1572
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1573	GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCACCAGGGCCACTGG	1632
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1633	GAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG	1692
Qy	1906	GCCAGTGCCTGCTACGTCTTACCCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1693	CCCAGGGCCTGTACATCCTGCTGGACCAGCTGGGCACCTACGTGTTACGGGCGAGTCC	1752

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:11:26 ; Search time 11689.8 Seconds
(without alignments)
11407.261 Million cell updates/sec

Title: US-10-624-932-1
Perfect score: 2752
Sequence: 1 ccgcggggcccccgcgcccgg.....tgagtgcctgaggccggccag 2752

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2752	100.0	2752	6	AX449572	AX449572 Sequence
2	2687.4	97.7	2697	6	AX451652	AX451652 Sequence
3	2676.4	97.3	2881	6	AX527916	AX527916 Sequence

4	2435.4	88.5	2784	6	CQ730306	CQ730306 Sequence
5	2356	85.6	3580	6	AX367094	AX367094 Sequence
6	2323.2	84.4	3992	10	MMU487852	AJ487852 Mus muscu
7	2259	82.1	3014	6	BD057524	BD057524 Netrin re
8	2252.2	81.8	2697	6	AX268596	AX268596 Sequence
9	2252.2	81.8	2697	10	RNU87305	U87305 Rattus norv
10	1983.6	72.1	3844	10	BC058084	BC058084 Mus muscu
11	1635.4	59.4	2688	9	BC009333	BC009333 Homo sapi
12	1562.4	56.8	1787	6	BD057525	BD057525 Netrin re
13	1304.2	47.4	9700	6	AX054976	AX054976 Sequence
14	992	36.0	9299	10	MMU72634	U72634 Mus musculu
15	991	36.0	2962	5	AY187310	AY187310 Gallus ga
16	986	35.8	9328	10	AB118026	AB118026 Rattus no
17	970	35.2	3646	6	CQ881060	CQ881060 Sequence
18	970	35.2	3646	9	AF055634	AF055634 Homo sapi
19	943.8	34.3	3770	9	AY126437	AY126437 Homo sapi
20	936.2	34.0	2860	6	AX686445	AX686445 Sequence
21	936.2	34.0	2860	6	AX686447	AX686447 Sequence
22	913.8	33.2	3672	10	BC048162	BC048162 Mus muscu
23	913.8	33.2	3672	10	BC057560	BC057560 Mus muscu
24	913.6	33.2	2895	6	AX512281	AX512281 Sequence
25	905.6	32.9	3933	6	CQ850929	CQ850929 Sequence
26	905.6	32.9	3933	9	AK128132	AK128132 Homo sapi
27	905.6	32.9	3935	9	AB096256	AB096256 Homo sapi
28	904.2	32.9	2986	6	CQ881052	CQ881052 Sequence
29	902.4	32.8	3884	6	AR528525	AR528525 Sequence
30	902.4	32.8	3884	6	AX464012	AX464012 Sequence
31	902.4	32.8	3884	9	AY358351	AY358351 Homo sapi
32	893.8	32.5	2995	6	AX497288	AX497288 Sequence
33	892	32.4	4294	10	AK122575	AK122575 Mus muscu
34	875.6	31.8	3788	10	MMU487853	AJ487853 Mus muscu
35	852.4	31.0	2838	10	RNU87306	U87306 Rattus norv
36	843.6	30.7	2625	6	CQ721377	CQ721377 Sequence
37	841.4	30.6	2831	6	BD057526	BD057526 Netrin re
38	764.8	27.8	813	6	AX054892	AX054892 Sequence
39	753.6	27.4	2230	6	CQ845766	CQ845766 Sequence
40	753.6	27.4	2230	9	AK131380	AK131380 Homo sapi
41	751.4	27.3	4330	5	AY744919	AY744919 Petromyzo
42	717.2	26.1	2832	5	AY099459	AY099459 Xenopus l
43	668.4	24.3	2612	6	CQ881064	CQ881064 Sequence
44	623.2	22.6	2661	6	AX800717	AX800717 Sequence
45	623.2	22.6	2868	6	AX800719	AX800719 Sequence

ALIGNMENTS

RESULT 1

AX449572

LOCUS AX449572 2752 bp DNA linear PAT 03-JUL-2002

DEFINITION Sequence 1 from Patent WO0210216.

ACCESSION AX449572

VERSION AX449572.1 GI:21698195

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 * AUTHORS Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
 Grosse,W.M., Alsobrook,J.P., Lepley,D.M., Gerlach,V.L.,
 Macdougall,J.R. and Smithson,G.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 0210216-A 1 07-FEB-2002;
 Curagen Corporation (US)
 FEATURES Location/Qualifiers
 source 1. .2752
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2752; DB 6; Length 2752;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	60
Db	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	60
Qy	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Db	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Qy	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Qy	181	TTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	181	TTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC	360
Db	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC	360
Qy	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGCTGGAGGAATACTGGTGCCAG	420
Db	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGCTGGAGGAATACTGGTGCCAG	420
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600

Db	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Qy	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACCTACACCTGCGTGGCCAAGAAC	720
Db	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACCTACACCTGCGTGGCCAAGAAC	720
Qy	721	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Db	721	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Qy	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGGAGC	960
Db	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGGAGC	960
Qy	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Db	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Qy	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Db	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Qy	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Db	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Qy	1141	ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Db	1141	ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Qy	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTACCTCAGGCTTCCAG	1260
Db	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTACCTCAGGCTTCCAG	1260
Qy	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCATCTGCTCACCATCCAGCCGGAC	1320
Db	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCATCTGCTCACCATCCAGCCGGAC	1320
Qy	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Db	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Qy	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440
Db	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440

Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Db	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Qy	1501	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Db	1501	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Qy	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Db	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Qy	1621	ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Db	1621	ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Qy	1681	TTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCCTGGC	1740
Db	1681	TTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCCTGGC	1740
Qy	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Db	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Qy	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Db	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Qy	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Db	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Qy	1921	GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980
Db	1921	GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980
Qy	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Db	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Qy	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Db	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Qy	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC	2160
Db	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC	2160
Qy	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Db	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Qy	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC	2280
Db	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC	2280

Qy 2281 TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG 2340
 |||
 Db 2281 TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG 2340

Qy 2341 GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC 2400
 |||
 Db 2341 GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC 2400

Qy 2401 ACAAGGTTTGTGTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC 2460
 |||
 Db 2401 ACAAGGTTTGTGTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC 2460

Qy 2461 CCCAGTGCCTTCAAGATCCCCTTCTCATTGGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
 |||
 Db 2461 CCCAGTGCCTTCAAGATCCCCTTCTCATTGGGCAGAAGATAATTTCCAGCCTGGACCCA 2520

Qy 2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCAT 2580
 |||
 Db 2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCAT 2580

Qy 2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
 |||
 Db 2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640

Qy 2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
 |||
 Db 2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700

Qy 2701 CCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
 |||
 Db 2701 CCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752

RESULT 2

AX451652

LOCUS AX451652 2697 bp DNA linear PAT 03-JUL-2002

DEFINITION Sequence 1 from Patent WO0233080.

ACCESSION AX451652

VERSION AX451652.1 GI:21698587

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Koehler, R.H.

TITLE Regulation of human netrin binding membrane receptor unc5h-1

JOURNAL Patent: WO 0233080-A 1 25-APR-2002;

Bayer Aktiengesellschaft (DE)

FEATURES

Location/Qualifiers

source

1. .2697

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

97.7%; Score 2687.4; DB 6; Length 2697;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGGCTC 105
      |||
Db      1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGGCTC 60

Qy     106 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC5CG 165
      |||
Db      61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy     166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG7CA 225
      |||
Db     121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy     226 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
      |||
Db     181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy     286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345
      |||
Db     241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy     346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTTCGGGCTGGAG 405
      |||
Db     301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTTCGGGCTGGAG 360

Qy     406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465
      |||
Db     361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy     466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525
      |||
Db     421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy     526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585
      |||
Db     481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Qy     586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645
      |||
Db     541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600

Qy     646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705
      |||
Db     601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660

Qy     706 TCGGTGGCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765
      |||
Db     661 TCGGTGGCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720

Qy     766 GTGAACGGTGGGTGGTGCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825
      |||
Db     721 GTGAACGGTGGGTGGTGCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780

Qy     826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGGCGCTTTC 885
      |||
```

Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Qy	886	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	945
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGC	900
Qy	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Db	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Qy	1006	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Db	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Qy	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC	1080
Qy	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATC	1185
Db	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATC	1140
Qy	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Db	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Qy	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Db	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Qy	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Db	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
Qy	1366	CAGGATGGGCCCAGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
Db	1321	CAGGATGGGCCCAGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Qy	1426	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485
Db	1381	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Qy	1486	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC	1560
Qy	1606	ATCCCCCAGATGCCATAACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1561	ATCCCCCAGATGCCATAACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Qy	1666	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGC	1680

Qy	1726	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1906	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1921	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2145
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2100
Qy	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2205
Db	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Qy	2206	AAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2265
Db	2161	AAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Qy	2266	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG	2325
Db	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG	2280
Qy	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
Db	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Qy	2386	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2445
Db	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2400
Qy	2446	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGCGCAGAAGATAATT	2505
Db	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGCGCAGAAGATAATT	2460
Qy	2506	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCAGAACTC	2565
Db	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCAGAACTC	2520

```

Qy      2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
          |||
Db      2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580

Qy      2626 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685
          |||
Db      2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640

Qy      2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGA 2742
          |||
Db      2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGA 2697

```

RESULT 3

AX527916

LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002

DEFINITION Sequence 1 from Patent WO0229038.

ACCESSION AX527916

VERSION AX527916.1 GI:25172359

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Herrmann, J.L., Rastelli, L. and Shimkets, R.A.

TITLE Novel proteins and nucleic acids encoding same and antibodies
directed against these proteins

JOURNAL Patent: WO 0229038-A 1 11-APR-2002;
Curagen Corporation (US)

FEATURES

source

Location/Qualifiers

1..2881

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 97.3%; Score 2676.4; DB 6; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

```

Qy      1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGGCCATGGCCGTCCGGCCC 60
          |||
Db      42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGGCCATGGCCGTCCGGCCC 101

Qy      61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
          |||
Db      102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161

Qy      121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
          |||
Db      162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy      181 TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
          |||
Db      222 TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281

```


Qy 241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
 |||
 Db 282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
 |||
 Qy 301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
 ||| |||
 Db 342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC 401
 |||
 Qy 361 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 420
 |||
 Db 402 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 461
 |||
 Qy 421 TGGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
 |||
 Db 462 TGGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521
 |||
 Qy 481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
 |||
 Db 522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581
 |||
 Qy 541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
 |||
 Db 582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
 |||
 Qy 601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660
 |||
 Db 642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701
 |||
 Qy 661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720
 |||
 Db 702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761
 |||
 Qy 721 ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTCATCGTCTACGTGAACGGTGGGTGG 780
 |||
 Db 762 ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTCATCGTCTACGTGAACGGTGGGTGG 821
 |||
 Qy 781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840
 |||
 Db 822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 881
 |||
 Qy 841 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900
 |||
 Db 882 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941
 |||
 Qy 901 GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG 957
 ||| | || || || || || ||
 Db 942 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 1001
 |||
 Qy 958 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1017
 |||
 Db 1002 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1061
 |||
 Qy 1018 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077
 |||
 Db 1062 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1121
 |||
 Qy 1078 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137

Db	1122	 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1197
Db	1182	 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	 AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	 ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATAACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	 GCCATAACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	 AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	 CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974

Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTTCGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTTCGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635	GAGGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679	GAGGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2796

LOCUS CQ730306 2784 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 16240 from Patent WO02068579.
 ACCESSION CQ730306
 VERSION CQ730306.1 GI:42303801
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 16240 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES Location/Qualifiers
 source 1. .2784
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 88.5%; Score 2435.4; DB 6; Length 2784;
 Best Local Similarity 93.5%; Pred. No. 0;
 Matches 2646; Conservative 0; Mismatches 6; Indels 177; Gaps 3;

Qy	46	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	105
Db	1	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	60
Qy	106	CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	165
Db	61	CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120
Qy	166	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	225
Db	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	180
Qy	226	GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	285
Db	181	GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Qy	286	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	345
Db	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Qy	346	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAAGGTGTTCTGGGCTGGAG	405
Db	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAAGGTGTTCTGGGCTGGAG	360
Qy	406	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	465
Db	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Qy	466	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	525

Db	421	TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Qy	526	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	585
Db	481	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Qy	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	645
Db	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Qy	646	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	705
Db	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Qy	706	TGCGTGGCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTTCATCGTCTAC	765
Db	661	TGCGTGGCCAAGAACATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTTCATCGTCTAC	720
Qy	766	-----	765
Db	721	GGTGGGCCCCGGGACTCCCTGGTCACAGGGAGAGGCACTGCGGTGCCCTGGGCAGTGAC	780
Qy	766	-----GTGAACGGTGGGTGGTTCGACGTGGACCGAGTGG	798
Db	781	ATGTGGCTGTCCTTCTCTGTCCGCCAGTGAACGGTGGGTGGTTCGACGTGGACCGAGTGG	840
Qy	799	TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC	858
Db	841	TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC	900
Qy	859	CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC	918
Db	901	CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC	960
Qy	919	GCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCCGGCCTGTGGG	978
Db	961	GCCACCCTGTGCCCAGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCCGGCCTGTGGG	1020
Qy	979	CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG	1038
Db	1021	CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG	1080
Qy	1039	GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC	1098
Db	1081	GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC	1140
Qy	1099	A-----GTGCTTCTGGCCCT	1113
Db	1141	AACTCCTACACCCCTGCCCCACCAAGGCCATGCTGTCTCCCGCAGCTGCTTCTGGCCCT	1200
Qy	1114	GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTG	1173
Db	1201	GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTG	1260
Qy	1174	CTTGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC	1233
Db	1261	CTTGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC	1320

Qy	1234	TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC	1293
Db	1321	TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC	1380
Qy	1294	CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGT	1353
Db	1381	CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGT	1440
Qy	1354	CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC	1413
Db	1441	CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC	1500
Qy	1414	AGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAG	1473
Db	1501	AGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAG	1560
Qy	1474	GAGTTCGTCTCCCGCCTCTCCACCCAGAATACTTCCGCTCCCTGCCCCGAGGCACCAGC	1533
Db	1561	GAGTTCGTCTCCCGCCTCTCCACCCAGAATACTTCCGCTCCCTGCCCCGAGGCACCAGC	1620
Qy	1534	AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGT	1593
Db	1621	AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGA	1680
Qy	1594	ATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC	1653
Db	1681	ATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC	1740
Qy	1654	ACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGT	1713
Db	1741	ACGCTGCACAAGCCGGAAGACGTG-----	1764
Qy	1714	CCCATCGTTAGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG	1773
Db	1765	-----AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG	1815
Qy	1774	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC	1833
Db	1816	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC	1875
Qy	1834	GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTAC	1893
Db	1876	GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTAC	1935
Qy	1894	TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGCTTTGCCCTG	1953
Db	1936	TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGCTTTGCCCTG	1995
Qy	1954	GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG	2013
Db	1996	GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG	2055
Qy	2014	GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGAT	2073
Db	2056	GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGAT	2115

Qy	2074	GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCA	2133
Db	2116	GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCA	2175
Qy	2134	CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCC	2193
Db	2176	CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCC	2235
Qy	2194	AGCTCCCTGTGGAAGAGTAAGCTCCTTGTGTCAGCTACCAGGAGATCCCCTTTTATCACATC	2253
Db	2236	AGCTCCCTGTGGAAGAGTAAGCTCCTTGTGTCAGCTACCAGGAGATCCCCTTTTATCACATC	2295
Qy	2254	TGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGC	2313
Db	2296	TGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGC	2355
Qy	2314	ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC	2373
Db	2356	ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC	2415
Qy	2374	AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT	2433
Db	2416	AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT	2475
Qy	2434	GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG	2493
Db	2476	GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG	2535
Qy	2494	CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG	2553
Db	2536	CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG	2595
Qy	2554	GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA	2613
Db	2596	GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA	2655
Qy	2614	GCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTG	2673
Db	2656	GCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTG	2715
Qy	2674	GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCT	2733
Db	2716	GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCT	2775
Qy	2734	GAGTGCTGA	2742
Db	2776	GAGTGCTGA	2784

RESULT 5

AX367094

LOCUS AX367094 3580 bp DNA linear PAT 16-FEB-2002

DEFINITION Sequence 13 from Patent WO0198354.

ACCESSION AX367094

VERSION AX367094.1 GI:18855296

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
Sanjanwala,M.S.

TITLE Receptors

JOURNAL Patent: WO 0198354-A 13 27-DEC-2001;
Incyte Genomics, Inc. (US)

FEATURES Location/Qualifiers
source 1. 3580
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 6052371CB1"

ORIGIN

Query Match 85.6%; Score 2356; DB 6; Length 3580;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2537; Conservative 0; Mismatches 5; Indels 168; Gaps 1;

Qy 43 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 102
|||||

Db 1 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60

Qy 103 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 162
|||||

Db 61 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120

Qy 163 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAG 222
|||||

Db 121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAG 180

Qy 223 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 282
|||||

Db 181 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240

Qy 283 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 342
|||||

Db 241 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 300

Qy 343 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTG 402
|||||

Db 301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTG 360

Qy 403 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 462
|||||

Db 361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420

Qy 463 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 522
|||||

Db 421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480

Qy 523 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 582

Db	481	GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGAGGGCATCCCTCCAGCC	540
Qy	583	GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC	642
Db	541	GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC	600
Qy	643	ATCACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTAC	702
Db	601	ATCACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTAC	660
Qy	703	ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTTCATCGTC	762
Db	661	ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTTCATCGTC	720
Qy	763	TACGTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG	822
Db	721	TACGTG-----	726
Qy	823	CGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGGCGCT	882
Db	727	-----	726
Qy	883	TTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	942
Db	727	-----GACGGC	732
Qy	943	AGCTGGAGCCCCTGGAGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1002
Db	733	AGCTGGAGCCCCTGGAGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	792
Qy	1003	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1062
Db	793	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	852
Qy	1063	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1122
Db	853	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG	912
Qy	1123	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTC	1182
Db	913	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTC	972
Qy	1183	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1242
Db	973	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1032
Qy	1243	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1302
Db	1033	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1092
Qy	1303	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1362
Db	1093	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1152
Qy	1363	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1422

Db	1153	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1212
Qy	1423	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC	1482
Db	1213	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC	1272
Qy	1483	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1542
Db	1273	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1332
Qy	1543	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1602
Db	1333	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC	1392
Qy	1603	CTCATCCCCCAGATGCCATACCCCAGGGGAAGATCTATGAGATCTACCTCAGCTGCAC	1662
Db	1393	CTCATCCCCCAGATGCCATACCCCAGGGGAAGATCTATGAGATCTACCTCAGCTGCAC	1452
Qy	1663	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1722
Db	1453	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1512
Qy	1723	AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1782
Db	1513	AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1572
Qy	1783	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1842
Db	1573	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1632
Qy	1843	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTG	1902
Db	1633	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTG	1692
Qy	1903	GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG	1962
Db	1693	GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG	1752
Qy	1963	GCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTTCGCGCCGGTGGCCTGC	2022
Db	1753	GCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTTCGCGCCGGTGGCCTGC	1812
Qy	2023	ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG	2082
Db	1813	ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG	1872
Qy	2083	GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG	2142
Db	1873	GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG	1932
Qy	2143	CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTG	2202
Db	1933	CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTG	1992
Qy	2203	TGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC	2262
Db	1993	TGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC	2052

Qy 2263 ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGAC 2322
 |||
 Db 2053 ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGAC 2112
 Qy 2323 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2382
 |||
 Db 2113 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2172
 Qy 2383 TTCAACATCACCAAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2442
 |||
 Db 2173 TTCAACATCACCAAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2232
 Qy 2443 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTTCGGCAGAAGATA 2502
 |||
 Db 2233 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTTCGGCAGAAGATA 2292
 Qy 2503 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2562
 |||
 Db 2293 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2352
 Qy 2563 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2622
 |||
 Db 2353 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2412
 Qy 2623 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2682
 |||
 Db 2413 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2472
 Qy 2683 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGA 2742
 |||
 Db 2473 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGA 2532
 Qy 2743 GGCCGGCCAG 2752
 |||
 Db 2533 GGCCGGCCAG 2542

RESULT 6

MMU487852

LOCUS MMU487852 3992 bp mRNA linear ROD 24-SEP-2002

DEFINITION Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).

ACCESSION AJ487852

VERSION AJ487852.1 GI:22035783

KEYWORDS netrin receptor Unc5h1; Unc5h1 gene.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Engelkamp, D.

TITLE Cloning of three mouse Unc5 genes and their expression patterns at
 mid-gestation

JOURNAL Mech. Dev. 118 (1-2), 191-197 (2002)

MEDLINE 22239710

PUBMED 12351186

REFERENCE 2 (bases 1 to 3992)

Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	427	AAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGGAATGGGTTCGCCAGGTC	486
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCCACCATGGAGGTCCGC	360
Db	487	GATCACGTCAATCGAACGCAGCACTGACGGCAGCAGCGGATTGCCAACCATGGAGGTCCGG	546
Qy	361	ATTAATGTCTCAAGGCAGCAGGTTCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	547	ATCAACGTATCAAGGCAGCAGGTTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAG	606
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	607	TGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCC	666
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	667	TATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTGCTACTGGAGCAAGGC	726
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	727	ATTGTGCTACCTTGTCGCCCCCGGAAGGAATCCCCCAGCTGAGGTGGAGTGGCTCCGA	786
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	787	AATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGC	846
Qy	661	CTGGTGGTGCAGCAGGCCCCGCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	847	CTAGTCGTGCGGCAGGCCCCGCTGGCCGACACTGCCAACTACACCTGCGTGGCCAAGAAC	906
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Db	907	ATCGTGGCCCCGTGCGCGAAGCGCCTCTGCGGCCGTGATTGTTTATGTGAACGGTGGGTGG	966
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	967	TCGACGTGGACCGAGTGGTCCGTCTGCAGTGCCAGCTGTGGGCGTGGCTGGCAGAAACGG	1026
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCGAAT	900
Db	1027	AGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTCTGTGAGGGGCGAAT	1086
Qy	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGGAGC	960
Db	1087	GTCCAGAAAACAGCCTGCGCCACTCTGTGCCAGTGGATGGGAGCTGGAGCCCATGGAGT	1146
Qy	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Db	1147	AAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGCCGGGAGTGCTCCGACCCA	1206
Qy	1021	GCACCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Db	1207	GCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGACACCCGCAACTGTACC	1266
Qy	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCCTC	1140

Db	1267	AGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTGGCTCTCTACATCGGCCTC	1326
Qy	1141	ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Db	1327	GTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCTCGTCCTCATCTACTGCCGC	1386
Qy	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG	1260
Db	1387	AAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATCCTTACCTCAGGCTTCCAG	1446
Qy	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC	1320
Db	1447	CCTGTCTAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTCACCATCCAACCGGAC	1506
Qy	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Db	1507	CTCAGCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGC	1566
Qy	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440
Db	1567	CCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACG	1626
Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAG	1500
Db	1627	CTGCACCACAGCTCCCCACCTCTGAGGCTGAGGACTTCGTCTCCCGCTCTCCACCCAA	1686
Qy	1501	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Db	1687	AACTACTTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCCTATGGGACCTTCAACTTC	1746
Qy	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Db	1747	CTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTCATACCCCCGGACGCC	1806
Qy	1621	ATACCCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Db	1807	ATACCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAGCCAGAAGACGTGAGG	1866
Qy	1681	TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC	1740
Db	1867	TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCATCGTTAGCTGTGGGCCCCCAGGA	1926
Qy	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Db	1927	GTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGAC	1986
Qy	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Db	1987	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGGGAGGACGTGCTGCAC	2046
Qy	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Db	2047	CTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCCTGCTAT	2106
Qy	1921	GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980

LOCUS BD057524 3014 bp DNA linear PAT 27-AUG-2002
 DEFINITION Netrin receptors.
 ACCESSION BD057524
 VERSION BD057524.1 GI:22603130
 KEYWORDS JP 2001505062-A/1.
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.
 REFERENCE 1 (bases 1 to 3014)
 AUTHORS Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
 TITLE Netrin receptors
 JOURNAL Patent: JP 2001505062-A 1 17-APR-2001;
 THE REGENTS OF THE UNIV OF CALIFORNIA
 COMMENT PN JP 2001505062-A/1
 PD 17-APR-2001
 PF 19-FEB-1998 JP 1998536840
 PR 19-FEB-1997 US 08/808982
 PI MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
 PI MASU,
 PI KAZUKO KEINO MASU
 PC C07K1/00, C07K14/00, C07K17/00, C07H21/02, C07H21/04, G01N33/53 CC
 Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES Location/Qualifiers
 source 1. .3014
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

ORIGIN

Query Match 82.1%; Score 2259; DB 6; Length 3014;
 Best Local Similarity 89.7%; Pred. No. 0;
 Matches 2427; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

Qy	46	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	105
Db	1	ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCTTGGCTT	60
Qy	106	CGCGGCTCGGGTGCCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	165
Db	61	CGTGGTTTCGGGTGCCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC	120
Qy	166	GACCTGCTTCCCCACTTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	225
Db	121	GACCTGCTGCCCCACTTCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG	180
Qy	226	GTGCTGCTTGTGTGCAAGGCCGTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	285
Db	181	GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA	240
Qy	286	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	345
Db	241	TGGGTCCGCCAGGTGCATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA	300
Qy	346	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAG	405

Db 301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTGTTGGGCTGGAG 360
 Qy 406 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465
 |||||
 Db 361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420
 Qy 466 TACATCCGCATAGCCAGATTGCGCAAGAAGTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525
 |||||
 Db 421 TACATCCGGATTGCCTATTTGCGCAAGAAGTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480
 Qy 526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585
 ||
 Db 481 TCACTGGAGCAAGGCATTGTACTACCTTGTGCCCCCAGAAGGAATCCCCCAGCTGAG 540
 Qy 586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645
 |||||
 Db 541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600
 Qy 646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705
 |||||
 Db 601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660
 Qy 706 TCGTGGCCAAGAATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCATCGTCTAC 765
 ||
 Db 661 TGTGTGGCCAAGAATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTATTGTTTAT 720
 Qy 766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825
 |||||
 Db 721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780
 Qy 826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885
 |||||
 Db 781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840
 Qy 886 TGTGAGGGGCGAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGC 945
 |||||
 Db 841 TGTGAGGGGCGAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCAGTGGATGGGAGC 900
 Qy 946 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 1005
 |||||
 Db 901 TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC 960
 Qy 1006 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1065
 |||||
 Db 961 GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGCGGGTGCTGACCTGGAC 1020
 Qy 1066 ACCCGCAACTGTACCAAGTACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1125
 |||||
 Db 1021 ACCCGCAACTGTACCAAGTACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT 1080
 Qy 1126 CTCTATGTGGGCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATC 1185
 |||||
 Db 1081 CTCTACATCGGCCTTGTGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA 1140
 Qy 1186 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1245
 |||||
 Db 1141 CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC 1200

Qy	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Db	1201	ACCTCGGGCTTCCAGCCTGTGAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qy	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Db	1261	ACCATCCAGCCAGACCTCAGCACCACCCTACCACCTACCAGGGCAGTCTATGTTGAGG	1320
Qy	1366	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
Db	1321	CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCCACTGGGG	1380
Qy	1426	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485
Db	1381	AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440
Qy	1486	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1441	CGCCTCTCCACCCAAAATACTTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC	1500
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC	1560
Qy	1606	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1561	ATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qy	1666	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1621	CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC	1680
Qy	1726	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1681	TGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA	1740
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1741	GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1801	GAGGATGTGCTGCACCTGGTGAGGAGTCACCTTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1906	GCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1861	GCCGGGGCCTGCTATGTCTTACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC	1920
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1921	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG	1980
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1981	TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG	2040

Qy 2086 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC 2145
 |||||
 Db 2041 GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCTGCAC 2100
 Qy 2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205
 |||||
 Db 2101 TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCAGCTCCCTGTGG 2160
 Qy 2206 AAGAGTAAGCTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCAGC 2265
 |||||
 Db 2161 AAGAGCAAGCTACTTGTCTAGCTACCAGGAGATCCCCTTTTACCACATCTGGAACGGCACC 2220
 Qy 2266 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG 2325
 |||||
 Db 2221 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280
 Qy 2326 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385
 |||||
 Db 2281 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340
 Qy 2386 AACATCACCAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGGTC 2445
 |||||
 Db 2341 AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400
 Qy 2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505
 |||||
 Db 2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460
 Qy 2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTC 2565
 |||||
 Db 2461 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAACTT 2520
 Qy 2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
 |||||
 Db 2521 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580
 Qy 2626 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685
 |||||
 Db 2581 AACCTATGGGAGGCGCGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640
 Qy 2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGC 2745
 |||||
 Db 2641 GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTGCGGAGGCCGAGTGTGAGAC 2700
 Qy 2746 CGGCCAG 2752
 |||||
 Db 2701 CAGCCAG 2707

RESULT 8

AX268596

LOCUS AX268596 2697 bp DNA linear PAT 29-OCT-2001

DEFINITION Sequence 15 from Patent WO0175440.

ACCESSION AX268596

VERSION AX268596.1 GI:16541710

KEYWORDS .

SOURCE Rattus sp.

ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1

AUTHORS Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and
Pratt,J.Y.

TITLE Schizophrenia related genes

JOURNAL Patent: WO 0175440-A 15 11-OCT-2001;
WELFIDE CORPORATION (JP)

FEATURES Location/Qualifiers

source 1. .2697
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN

Query Match 81.8%; Score 2252.2; DB 6; Length 2697;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

Qy	46	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	105
Db	1	ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGTCTCCTGGGCATAGTCCTCGCCGCTTGGCTT	60
Qy	106	CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	165
Db	61	CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC	120
Qy	166	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	225
Db	121	GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG	180
Qy	226	GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	285
Db	181	GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA	240
Qy	286	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	345
Db	241	TGGGTCCGCCAGGTGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA	300
Qy	346	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGGGGCTGGAG	405
Db	301	ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG	360
Qy	406	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	465
Db	361	GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC	420
Qy	466	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	525
Db	421	TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG	480
Qy	526	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	585
Db	481	TCACTGGAGCAAGGCATTGTACTACCTTGTGCCCCCAGAAGGAATCCCCCAGCTGAG	540

Qy	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	645
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
Qy	646	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACC	705
Db	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC	660
Qy	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGACGCGCTCCGCTGCTGTCATCGTCTAC	765
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCAATTGTTTAT	720
Qy	766	GTGAACGGTGGGTGGTGCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	825
Db	721	GTGAACGGTGGGTGGTGCACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT	780
Qy	826	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	885
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC	840
Qy	886	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	945
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	900
Qy	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qy	1006	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGCTGACCTGGAC	1020
Qy	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qy	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATC	1185
Db	1081	CTCTACATCGGCCTTGTGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA	1140
Qy	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Db	1141	CTCATTTACTGTGCGAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC	1200
Qy	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Db	1201	ACCTCGGGCTTCCAGCCTGTGAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qy	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Db	1261	ACCATCCAGCCAGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTATGTTGAGG	1320
Qy	1366	CAGGATGGGCCCAGCCCCAAGTTCAGCTACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
Db	1321	CAGGATGGACCCAGCCCCAAGTTCAGCTCTCTAATGGTCACCTGCTCAGCCCCACTGGGG	1380
Qy	1426	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485

Db	1381	AGTGGCCGCCATACGTTGCACCACAGCTACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440
Qy	1486	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1441	CGCCTCTCCACCCAAACTACTTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC	1500
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC	1560
Qy	1606	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1561	ATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qy	1666	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1621	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC	1680
Qy	1726	TGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1681	TGTGGGCCCCCAGGAGTCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA	1740
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1741	GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1801	GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1906	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCGCCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1861	GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCGCCTTTGCCCTGGTAGGAGAGGCC	1920
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1921	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG	1980
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1981	TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG	2040
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2145
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCTGCAC	2100
Qy	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205
Db	2101	TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCAGCTCCCTGTGG	2160
Qy	2206	AAGAGTAAGCTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACC	2265
Db	2161	AAGAGCAAGCTACTTGTGAGCTACCAGGAGATCCCTTTTACCACATCTGGAACGGCACC	2220
Qy	2266	CAGCGGTACTTGCAGTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG	2325

Db	2221	CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG	2280
Qy	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
Db	2281	GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC	2340
Qy	2386	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2445
Db	2341	AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC	2400
Qy	2446	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT	2505
Db	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC	2460
Qy	2506	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTC	2565
Db	2461	GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAACTT	2520
Qy	2566	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2625
Db	2521	CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC	2580
Qy	2626	AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2685
Db	2581	AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG	2640
Qy	2686	GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGA	2742
Db	2641	GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTGCGAGGCCGAGTGTTGA	2697

JOURNAL Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San Francisco, CA 94143-0452, USA

```

FEATURES                      Location/Qualifiers
    source                    1. .2697
                              /organism="Rattus norvegicus"
                              /mol_type="mRNA"
                              /db_xref="taxon:10116"
                              /tissue_type="brain and ventral spinal cord"
                              /dev_stage="18 day embyo and 13 day embryo"
    CDS                       1. .2697
                              /codon_start=1
                              /product="transmembrane receptor UNC5H1"
                              /protein_id="AAB57678.1"
                              /db_xref="GI:2055392"
                              /translation="MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLP
HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTM
EVRINVSRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
SLEQGIVLPCRPPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
YTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRCTNPAPLN
GGAFCGEQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWSRECSDPAPRNGGEEC
RGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS
DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGSLSRQDGPSPKFQL
SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRSLSTQNYFRSLPRGTSNMAYGTFFNLGG
RLMIPNTGISLLIPDAIPRGKIYEIYLTLLHKPEDVRLPLAGCQTLLSPVWSCGPPGV
LLTRPVILAMDHCGEPPSPDSWSLRLKKQSCGSWEDVLHLGEESPSHLYYCQLEAGAC
YVFTEQLGRFALVGEALSAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGT
QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDRFAELLALESEG
GVPALVGPSAFKIPFLIRQKIIASLDPPCSRGA DWRTLAQKLHLD SHLSFFASKPSPT
AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAE"

```

ORIGIN

```

Query Match          81.8%;  Score 2252.2;  DB 10;  Length 2697;
Best Local Similarity 89.7%;  Pred. No. 0;
Matches 2419;  Conservative 0;  Mismatches 278;  Indels 0;  Gaps 0;

```

```

Qy      46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
        ||||||||||||||||||||||||||| ||||||||||||||||||||||| |||||
Db      1  ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCTTGGCTT 60

Qy     106 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
        || || ||||||||||||||||||||||| ||||||| ||||||| |||||||||||
Db      61 CGTGGTTCGGGTGCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy     166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
        ||||||| ||||||||||||||| ||||| ||||| ||||||| |||||||||||
Db     121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy     226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
        ||| || | ||||||||||| ||||| ||||| ||||||||||| |||||
Db     181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy     286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345
        ||||| ||||||| || ||||| || || ||||||| ||| | ||||| || |||||
Db     241 TGGGTCCGCCAGGTGCATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy     346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAAGGTGTTGGGCTGGAG 405
        ||||||||||||||| || || || || ||||||||||| ||||| ||||| |||||||
Db     301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

```


Qy	406	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	465
Db	361	GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC	420
Qy	466	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	525
Db	421	TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG	480
Qy	526	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGAGGGGCATCCCTCCAGCCGAG	585
Db	481	TCCTGGAGCAAGGCATTGTACTACCTTGTGCCCCCAGAAGGAATCCCCCAGCTGAG	540
Qy	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	645
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
Qy	646	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACC	705
Db	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC	660
Qy	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGACAGCGCCTCCGCTGCTGTCATCGTCTAC	765
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT	720
Qy	766	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	825
Db	721	GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT	780
Qy	826	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	885
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC	840
Qy	886	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	945
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	900
Qy	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qy	1006	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC	1020
Qy	1066	ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
Db	1021	ACCCGCAACTGTACCAAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qy	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCCTGCTGCTGCTTGTCTCATC	1185
Db	1081	CTCTACATCGGCCTTGTGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGA	1140
Qy	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Db	1141	CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCTC	1200

[illegible]

```

      |||
Db      2041 GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC 2100
      |||
Qy      2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205
      |||
Db      2101 TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG 2160
      |||
Qy      2206 AAGAGTAAGCTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265
      |||
Db      2161 AAGAGCAAGCTACTTGTCTAGCTACCAGGAGATCCCCTTTTACCACATCTGGAACGGCACC 2220
      |||
Qy      2266 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG 2325
      |||
Db      2221 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280
      |||
Qy      2326 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385
      |||
Db      2281 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340
      |||
Qy      2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGGTC 2445
      |||
Db      2341 AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGGTC 2400
      |||
Qy      2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505
      |||
Db      2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460
      |||
Qy      2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTC 2565
      |||
Db      2461 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAACTT 2520
      |||
Qy      2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
      |||
Db      2521 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580
      |||
Qy      2626 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685
      |||
Db      2581 AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640
      |||
Qy      2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGA 2742
      |||
Db      2641 GCCGGAAGTGGGCCAACAGATGCTGGCCTCTTCACGGTGTGCGGAGGCCGAGTGTTGA 2697

```

RESULT 10

BC058084

LOCUS BC058084 3844 bp mRNA linear ROD 21-OCT-2003

DEFINITION Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone MGC:66671 IMAGE:6813463), complete cds.

ACCESSION BC058084

VERSION BC058084.1 GI:34784158

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3844)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 3844)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 126 Row: b Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23346570.

FEATURES Location/Qualifiers

source 1. .3844
/organism="Mus musculus"
/mol_type="mRNA"

```

/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:66671 IMAGE:6813463"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_lib="NIH_BMAP_FW0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
gene 1. .3844
      /gene="Unc5a"
      /db_xref="LocusID:107448"
      /db_xref="MGI:894682"
CDS 252. .2780
      /gene="Unc5a"
      /codon_start=1
      /product="Unc5a protein"
      /protein_id="AAH58084.1"
      /db_xref="GI:34784159"
      /db_xref="LocusID:107448"
      /db_xref="MGI:894682"
      /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
SLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTPN
YTCVAKNIVARRRSASAAVIVYVDGWSWPWSKWSACGLDCTHWSRECSDPAPRNGGE
ECRGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGL
DSDVADSSILTSFGQPVSISKPSKADNPHELLTIQPDLSSTTTTTYQGSCLPRQDGPSPKF
QLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFL
GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPP
GVLLTRPVILAMDHCGEPPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAG
ACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKE
VVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSCLLVSYQEIPFYHIWN
GTQQYLHCTFTLERVNASTD LACKVWVWQVEGDGQSFNINFNITKDRFAEMLALES
EGGVPALVGPSAFKIPFLIRQKIITS LDPPCSRGADWRTLAQKLHLD SHLSFFASKPS
PTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
misc_feature 714. .926
      /gene="Unc5a"
      /note="IG; Region: Immunoglobulin"
      /db_xref="CDD:smart00409"
misc_feature 984. .1130
      /gene="Unc5a"
      /note="TSP1; Region: Thrombospondin type 1 repeats"
      /db_xref="CDD:smart00209"
misc_feature 1566. .1877
      /gene="Unc5a"
      /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
netrin receptors"
      /db_xref="CDD:smart00218"
misc_feature 2496. .2684
      /gene="Unc5a"
      /note="DEATH; Region: DEATH domain, found in proteins
involved in cell death (apoptosis). Alpha-helical domain
present in a variety of proteins with apoptotic functions.
Some (but not all) of these domains form homotypic and
heterotypic dimers"
      /db_xref="CDD:smart00005"

```

ORIGIN

Query Match 72.1%; Score 1983.6; DB 10; Length 3844;
Best Local Similarity 84.3%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 264; Indels 168; Gaps 1;

Qy	1	CCGCGGGGCCCCCGCCGCCCCGCCCCGCTGCCGCCCGCGGCCATGGCCGTCGGGCC	60
Db	207	CCGCCCCGCGGGGCCCCCTAGTCCAGCCCGCTCGTCCGCCCGCGGCCATGGCCGTCGGGCC	266
Qy	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Db	267	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTTCGTGGTTCGGGTGCC	326
Qy	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	327	CAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTGCCCCAC	386
Qy	181	TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	387	TTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCCGTGCTGCTGGTGTGC	446
Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	447	AAGGCTGTGCCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAATGGGTTCGCCAGGTC	506
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC	360
Db	507	GATCACGTCAATTGAACGCAGCACTGACGGCAGCAGCGGATTGCCAACCATGGAGGTCCGG	566
Qy	361	ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	567	ATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAG	626
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	627	TGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCC	686
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	687	TATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTGTCCTGGAGCAAGGC	746
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	747	ATTGTGCTACCTTGTGCCCCCGGAAGGAATCCCCCAGCTGAGGTGGAGTGGTCCGA	806
Qy	601	AACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	807	AATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGC	866
Qy	661	CTGGTGGTGGCAGAGGCCCGCTTGTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	867	CTAGTCGTGCGGCAGGCCCGCTGGCCGACACTCCCAACTACACCTGCGTGGCCAAGAAC	926
Qy	721	ATCGTGGCACGTCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	780
Db	927	ATCGTGGCCCCGTGCGCGAAGCGCCTCTGCGGCCGTATTGTTTATGTG-----	974
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840

Db	975	-----	974
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	975	-----	974
Qy	901	GTCCAGAAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGGAGC	960
Db	975	-----GATGGGAGCTGGAGCCCATGGAGT	998
Qy	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Db	999	AAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGCCGGGAGTGCTCCGACCCA	1058
Qy	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Db	1059	GCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGACACCCGCAACTGTACC	1118
Qy	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Db	1119	AGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTGGCTCTCTACATCGGCCTC	1178
Qy	1141	ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGCCGG	1200
Db	1179	GTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCTCTCGTCTCATCTACTGCCGC	1238
Qy	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG	1260
Db	1239	AAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATCCTTACCTCAGGCTTCCAG	1298
Qy	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC	1320
Db	1299	CCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTCACCATCCAACCGGAC	1358
Qy	1321	CTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Db	1359	CTCAGCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGC	1418
Qy	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440
Db	1419	CCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACG	1478
Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Db	1479	CTGCACCACAGTCTCCCCACCTCTGAGGCTGAGGACTTCGTCTCCCGCCTCTCCACCCAA	1538
Qy	1501	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Db	1539	AACTACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCCTATGGGACCTTCAACTTC	1598
Qy	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Db	1599	CTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTCATACCCCCGGACGCC	1658
Qy	1621	ATACCCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680

Db	1659	ATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAGCCAGAAGACGTGAGG	1718
Qy	1681	TTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC	1740
Db	1719	TTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCATCGTTAGCTGTGGGCCCCCAGGA	1778
Qy	1741	GTCCTGCTCACCCGGCCAGTCATCCTTGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Db	1779	GTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGAC	1838
Qy	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Db	1839	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGGGAGGACGTGCTGCAC	1898
Qy	1861	CTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Db	1899	CTTGGTGAGGAGTGCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCTGCTAT	1958
Qy	1921	GTCTTCACCGAGCAGCTGGGCGCTTTGCCCTGGTGAGAGGCCCTCAGCGTGGCTGCC	1980
Db	1959	GTCTTTACCAGCAGCTAGGCCGCTTTGCCCTGGTGAGAGGCCCTCAGCGTGGCTGCC	2018
Qy	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Db	2019	ACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCTGTGGCCTGTACGTCCCTCGAGTACAAC	2078
Qy	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Db	2079	ATCCGAGTGTA CTGCTGCACGACACCCACGATGCTCTCAAGGAGGTGGTGCAGCTGGAG	2138
Qy	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTAC	2160
Db	2139	AAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCTGCACTTCAAAGACAGTTAC	2198
Qy	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Db	2199	CACAACCTACGTCTGTCCATCCACGACGTGCCAGCTCCCTGTGGAAGAGCAAGCTCCTT	2258
Qy	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGAC	2280
Db	2259	GTCAGCTACCAGGAGATCCCTTTTACCACATCTGGAATGGCACTCAGCAGTATCTGCAC	2318
Qy	2281	TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG	2340
Db	2319	TGCACCTTCACCCTGGAGCGGTCAATGCCAGCACCAGCGACCTGGCCTGCAAGGTGTGG	2378
Qy	2341	GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC	2400
Db	2379	GTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTTAACATCACTAAGGAC	2438
Qy	2401	ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC	2460
Db	2439	ACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTCCCAGCCCTGGTGGGC	2498
Qy	2461	CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA	2520
Db	2499	CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATTACCAGCCTGGACCCA	2558

Qy	2521	CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT	2580
Db	2559	CCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTTCACCTGGACAGCCAT	2618
Qy	2581	CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG	2640
Db	2619	CTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTCAACCTATGGGAGGCG	2678
Qy	2641	CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG	2700
Db	2679	CGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTGGCCGGACTGGGCCAG	2738
Qy	2701	CCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	CCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGCTGAGACCAGCCAG	2790

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:14424611.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: g Column: 22.

FEATURES

source

Location/Qualifiers
1. .2688
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4126760"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

gene

<1. .2688
/gene="UNC5A"
/note="synonym: UNC5H1"
/db_xref="LocusID:90249"
/db_xref="MIM:607869"

CDS

<1. .1627
/gene="UNC5A"
/codon_start=2
/product="UNC5A protein"
/protein_id="AAH09333.2"
/db_xref="GI:40226528"
/db_xref="LocusID:90249"
/db_xref="MIM:607869"
/translation="DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT
SGFQPVSIKPSKADNPHELLTIQPDLTSTTTTYYQGSCLPRQDGPSPKFQLTNGHLLSPL
GGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGI
SLIIPDAIPRGKIYEIYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILA
MDHCGEPPSPDSWSLRLKKQSCGSEWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR
FALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQ
LIQEPRVLHFKDSYHNLRSLSIHDVPSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFT

```

LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPS
AFKIPFLIRQKIISLDPPCRRGADWRTLQKLHLDShLSFFASKPSPTAMILNLWEA
RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"
misc_feature 413. .724
               /gene="UNC5A"
               /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
               netrins receptors"
               /db_xref="CDD:smart00218"
misc_feature 1343. .1594
               /gene="UNC5A"
               /note="DEATH; Region: DEATH domain, found in proteins
               involved in cell death (apoptosis). Alpha-helical domain
               present in a variety of proteins with apoptotic functions.
               Some (but not all) of these domains form homotypic and
               heterotypic dimers"
               /db_xref="CDD:smart00005"

```

ORIGIN

```

Query Match          59.4%; Score 1635.4; DB 9; Length 2688;
Best Local Similarity 99.9%; Pred. No. 2.4e-264;
Matches 1636; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1116 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 1175
          |||
Db      1   GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 60

Qy      1176 TGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 1235
          |||
Db      61   TGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 120

Qy      1236 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 1295
          |||
Db      121  GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 180

Qy      1296 CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 1355
          |||
Db      181  CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 240

Qy      1356 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 1415
          |||
Db      241  CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 300

Qy      1416 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGA 1475
          |||
Db      301  CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGA 360

Qy      1476 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 1535
          |||
Db      361  GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 420

Qy      1536 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTAT 1595
          |||
Db      421  CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAAT 480

Qy      1596 CAGCCTCCTCATCCCCCAGATGCCATAACCCGAGGGAAGATCTATGAGATCTACCTCAC 1655
          |||
Db      481  CAGCCTCCTCATCCCCCAGATGCCATAACCCGAGGGAAGATCTATGAGATCTACCTCAC 540

```

Qy	1656	GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC	1715
Db	541	GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC	600
Qy	1716	CATCGTTAGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA	1775
Db	601	CATCGTTAGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA	660
Qy	1776	CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA	1835
Db	661	CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA	720
Qy	1836	GGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTG	1895
Db	721	GGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTG	780
Qy	1896	CCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGT	1955
Db	781	CCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGT	840
Qy	1956	GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT	2015
Db	841	GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT	900
Qy	2016	GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGC	2075
Db	901	GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGC	960
Qy	2076	ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG	2135
Db	961	ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG	1020
Qy	2136	GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAG	2195
Db	1021	GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAG	1080
Qy	2196	CTCCCTGTGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTG	2255
Db	1081	CTCCCTGTGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTG	1140
Qy	2256	GAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCAC	2315
Db	1141	GAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCAC	1200
Qy	2316	TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAG	2375
Db	1201	TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAG	1260
Qy	2376	CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA	2435
Db	1261	CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA	1320
Qy	2436	AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA	2495
Db	1321	AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA	1380

```

Qy      2496 GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC 2555
          |||
Db      1381 GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC 1440

Qy      2556 CCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC 2615
          |||
Db      1441 CCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC 1500

Qy      2616 CATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGC 2675
          |||
Db      1501 CATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGC 1560

Qy      2676 TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGA 2735
          |||
Db      1561 TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGA 1620

Qy      2736 GTGCTGAGGCCGGCCAG 2752
          |||
Db      1621 GTGCTGAGGCCGGCCAG 1637

```

RESULT 12

BD057525

LOCUS BD057525 1787 bp DNA linear PAT 27-AUG-2002

DEFINITION Netrin receptors.

ACCESSION BD057525

VERSION BD057525.1 GI:22603131

KEYWORDS JP 2001505062-A/2.

SOURCE synthetic construct

ORGANISM synthetic construct

other sequences; artificial sequences.

REFERENCE 1 (bases 1 to 1787)

AUTHORS Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.

TITLE Netrin receptors

JOURNAL Patent: JP 2001505062-A 2 17-APR-2001;

THE REGENTS OF THE UNIV OF CALIFORNIA

COMMENT PN JP 2001505062-A/2

PD 17-APR-2001

PF 19-FEB-1998 JP 1998536840

PR 19-FEB-1997 US 08/808982

PI MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI

PI MASU,

PI KAZUKO KEINO MASU

PC C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC

Strandedness: Double;

CC Topology: Linear;

FH Key Location/Qualifiers.

FEATURES Location/Qualifiers

source

1. .1787

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

ORIGIN

Query Match 56.8%; Score 1562.4; DB 6; Length 1787;

Best Local Similarity 98.5%; Pred. No. 4.3e-252;

Matches 1661; Conservative 0; Mismatches 16; Indels 9; Gaps 8;

Qy	1070	GCAACTGTACCACTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT	1129
Db	1	GCAACTGTACCACTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT	59
Qy	1130	ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCCTCATCCTCG	1189
Db	60	ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCCTCATCCTCG	119
Qy	1190	TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT	1249
Db	120	TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT	179
Qy	1250	CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC	1308
Db	180	CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC	239
Qy	1309	ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG	1368
Db	240	ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG	299
Qy	1369	GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC	1428
Db	300	GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC	359
Qy	1429	GGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC	1488
Db	360	GGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC	419
Qy	1489	CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG	1548
Db	420	CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG	479
Qy	1549	ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC	1608
Db	480	ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC	539
Qy	1609	CCCCCAGATGCCATACCCCAGGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG	1668
Db	540	CCCCCAGATGCCATACCCCAGGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG	599
Qy	1669	GAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT	1728
Db	600	GAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT	659
Qy	1729	GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG	1788
Db	660	GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG	719
Qy	1789	CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG	1848
Db	720	CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG	778
Qy	1849	GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCC	1908
Db	779	GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCC	837

Qy 1909 AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC 1968
 |||
 Db 838 AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC 897

Qy 1969 AGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC 2028
 |||
 Db 898 AGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC 957

Qy 2029 CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG 2088
 |||
 Db 958 CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG 1017

Qy 2089 GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTC 2148
 |||
 Db 1018 GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTC- 1076

Qy 2149 AAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAG 2208
 ||| |||
 Db 1077 AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGGAAG 1134

Qy 2209 AGTAAGCTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG 2268
 |||
 Db 1135 AGTAAGCTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG 1194

Qy 2269 CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCC 2328
 |||
 Db 1195 CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCC 1254

Qy 2329 TGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC 2388
 |||
 Db 1255 TGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC 1314

Qy 2389 ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA 2448
 |||
 Db 1315 ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA 1374

Qy 2449 GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC 2508
 |||
 Db 1375 GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC 1434

Qy 2509 AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCAC 2568
 |||
 Db 1435 AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCAC 1494

Qy 2569 CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC 2628
 |||
 Db 1495 CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC 1554

Qy 2629 CTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCT 2688
 |||
 Db 1555 CTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCT 1614

Qy 2689 GGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAGGCC 2746
 || ||| ||| ||| |||
 Db 1615 GGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTTCGGAGGCTGAGTGCTGAGGCC 1674

Qy 2747 GGCCAG 2752

|||||
Db 1675 GGCCAG 1680

RESULT 13

AX054976

LOCUS AX054976 9700 bp DNA linear PAT 13-JAN-2001

DEFINITION Sequence 91 from Patent WO0073328.

ACCESSION AX054976

VERSION AX054976.1 GI:12228344

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct

other sequences; artificial sequences.

REFERENCE 1

AUTHORS van Criekinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.

TITLE Unc-5 constructs and screening methods

JOURNAL Patent: WO 0073328-A 91 07-DEC-2000;

Devgen NV (BE)

FEATURES Location/Qualifiers

source

1. .9700

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="plasmid pGC1037"

ORIGIN

Query Match 47.4%; Score 1304.2; DB 6; Length 9700;

Best Local Similarity 99.8%; Pred. No. 6.8e-209;

Matches 1306; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1437 CACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 1496

|||||

Db 7 CACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 66

Qy 1497 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 1556

|||||

Db 67 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 126

Qy 1557 CTTCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA 1616

|||||

Db 127 CTTCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCAGA 186

Qy 1617 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 1676

|||||

Db 187 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 246

Qy 1677 GAGGTTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 1736

|||||

Db 247 GAGGTTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 306

Qy 1737 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 1796

|||||

Db 307 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 366

Qy 1797 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT 1856

|||||

Db	367	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	426
Qy	1857	GCACCTGGGCGAGGAGGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	1916
Db	427	GCACCTGGGCGAGGAGGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	486
Qy	1917	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	1976
Db	487	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	546
Qy	1977	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA	2036
Db	547	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA	606
Qy	2037	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	2096
Db	607	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	666
Qy	2097	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG	2156
Db	667	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG	726
Qy	2157	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT	2216
Db	727	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT	786
Qy	2217	CCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT	2276
Db	787	CCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT	846
Qy	2277	GCACTGCACCTTCACCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT	2336
Db	847	GCACTGCACCTTCACCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT	906
Qy	2337	GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA	2396
Db	907	GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA	966
Qy	2397	GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT	2456
Db	967	GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCAAGCCCTGGT	1026
Qy	2457	GGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTGCGCAGAAGATAATTTCCAGCCTGGA	2516
Db	1027	GGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTGCGCAGAAGATAATTTCCAGCCTGGA	1086
Qy	2517	CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAG	2576
Db	1087	CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAG	1146
Qy	2577	CCATCTCAGCTTCTTTGCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA	2636
Db	1147	CCATCTCAGCTTCTTTGCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA	1206
Qy	2637	GGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG	2696
Db	1207	GGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG	1266

Qy 2697 CCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGC 2745
 |||||||||||||||||||||||||||||||||||||||||||||
 Db 1267 CCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGTC 1315

RESULT 14

MMU72634

LOCUS MMU72634 9299 bp mRNA linear ROD 13-MAY-1997

DEFINITION Mus musculus rostral cerebellar malformation protein (rcm) mRNA, complete cds.

ACCESSION U72634

VERSION U72634.1 GI:2088526

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 9299)

AUTHORS Ackerman, S.L., Kozak, L.P., Przyborski, S.A., Rund, L.A., Boyer, B.B. and Knowles, B.B.

TITLE The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein

JOURNAL Nature 386 (6627), 838-842 (1997)

MEDLINE 97271898

PUBMED 9126743

REFERENCE 2 (bases 1 to 9299)

AUTHORS Ackerman, S.L., Kozak, L.P., Rund, L.A. and Knowles, B.B.

TITLE Direct Submission

JOURNAL Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA

FEATURES Location/Qualifiers

source

1. .9299
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57B6/SJL"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="between Amy1 and Mpmv9"

gene

1. .9299
 /gene="rcm"

CDS

151. .2946
 /gene="rcm"
 /function="transmembrane receptor"
 /note="similar to C. elegans UNC-5"
 /codon_start=1
 /product="rostral cerebellar malformation protein"
 /protein_id="AAB54103.1"
 /db_xref="GI:2088527"
 /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE
 FFHELPETFPSPDPPEPLPHFLIEPEEAYIVKNKPNLYCKASPATQIYFCNSEWVHQ
 KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
 RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
 TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
 GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTSWSKWSCTGTECTH
 WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPSDDDVALYVGIVIAVTVCLA
 ITVVVALFVYRKNHRDFESDIIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR

Db	908	CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTGTGA	967
Qy	809	GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	868
Db	968	ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC	1027
Qy	869	TCAACGGGGGCGCTTTCTGTGAGGGGCGAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT	928
Db	1028	TCAATGGTGGGGCCTTCTGTGAGGGGCGAGAGTGTGCAGAAAATAGCATGCACTACGTTAT	1087
Qy	929	GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA	988
Db	1088	GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAATGCA	1147
Qy	989	CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC	1048
Db	1148	CCCACTGGCGCAGGAGGGAGTGACAGCACCAGCCCCCAAGAACGGGGGTAAGGACTGTG	1207
Qy	1049	AGGGCACTGACCTGGACACCCGCAACTGTACCACTGACCTCTGTGTACACAGTGCTTCTG	1108
Db	1208	ATGGCCTGGTCTCTCAATCCAAGAACTGCACTGATGGGCTGTGCATGCAGGCTGCTCCTG	1267
Qy	1109	GCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCC	1165
Db	1268	ACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCTGGCGA	1327
Qy	1166	TGCTGCTGCTTGTCTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATG	1225
Db	1328	TCACTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACA	1387
Qy	1226	TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG	1285
Db	1388	TCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAG-----G	1438
Qy	1286	CAGACAACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACC	1345
Db	1439	CTGCCAGACAAGATCTCCTGGCTGTCCCCCTGACCTCACCTCAGCTGCAGCCATGTACA	1498
Qy	1346	AGGGCAGTCTCTGTCCCCGGCAGGATG-----GGCCCAGCCCCAAGTTCAGCTCACCAA	1400
Db	1499	GGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCAA	1558
Qy	1401	TGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCAC	1460
Db	1559	TTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTC	1618
Qy	1461	CTC-----TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC-----	1498
Db	1619	CTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGC	1678
Qy	1499	-----AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA	1534
Db	1679	TAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGACTGACCCATCCT	1738
Qy	1535	ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA	1594
Db	1739	GCACAGCATTGGTACCTTCAACTCTCTGGGGGTACCTCATCATTCCTAATTCAGGAG	1798

Qy	1595	TCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA	1654
Db	1799	TAAGCTTGCTGATTCCCGCTGGGGCCATTCTCAGGGGAGAGTCTATGAAATGTATGTGA	1858
Qy	1655	CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTC	1714
Db	1859	CTGTACACAGGAAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCC	1918
Qy	1715	CCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGG	1774
Db	1919	CTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCTCCTCACTCTGC	1978
Qy	1775	ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG	1834
Db	1979	ATCACTGTGCAGACCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGC	2038
Qy	1835	AGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACT	1894
Db	2039	AGGGACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCTGTTACA	2098
Qy	1895	GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG	1954
Db	2099	TTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGG	2158
Qy	1955	TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG	2014
Db	2159	TTGGGCAGTCCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCC	2218
Qy	2015	TGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATG	2074
Db	2219	TCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATG	2278
Qy	2075	CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCAC	2134
Db	2279	CCCTGAAGGAAGTTCTACAACCTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCA	2338
Qy	2135	GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCA	2194
Db	2339	AGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTTCATGACATCGCCC	2398
Qy	2195	GCTCCCTGTGGAAGAGTAAGCTCCTTGTGCTAGCTACCAGGAGATCCCCCTTTTATCACATCT	2254
Db	2399	ATTCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCT	2458
Qy	2255	GGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCA	2314
Db	2459	GGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAGACTCAGCCTAAACA	2518
Qy	2315	CTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCA	2374
Db	2519	CAGTGGAAGTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCC	2578
Qy	2375	GCATCAACTTCAACATCACCAAGGACACAAGGTTTGTGAGCTGCTGGCTCTGGAGAGTG	2434
Db	2579	AGCTCAACTGTACTGTGTGAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTG	2638

Qy 2435 AAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGC 2494
 || | | | | | | | | | | | | | | | | | |
 Db 2639 CTAGTACCATCACCCTGTACCAGGACCAAGTGCTTTTCAGCATTCTCTCCCTATCCGGC 2698
 Qy 2495 AGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGG 2554
 || || | | | | | | | | | | | | | | | | | |
 Db 2699 AGAAGCTATGCAGCAGCCTGGATGCCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGG 2758
 Qy 2555 CCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAG 2614
 || | | | | | | | | | | | | | | | | | | | | | |
 Db 2759 CCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAACTG 2818
 Qy 2615 CCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGG 2674
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2819 GCGTAATCCTGGATCTTTGGGAAGCACAGAAGCTTCCCAGATGGAAACCTGAGCATGCTGG 2878
 Qy 2675 CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTG 2734
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2879 CAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCTTGGCAGCAGAAGGAC 2938
 Qy 2735 AGTGCTGA 2742
 || | || |
 Db 2939 AGTATTGA 2946

RESULT 15

AY187310

LOCUS AY187310 2962 bp mRNA linear VRT 06-JUN-2003

DEFINITION Gallus gallus UNC5-like protein 3 mRNA, complete cds.

ACCESSION AY187310

VERSION AY187310.1 GI:31442350

KEYWORDS .

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 2962)

AUTHORS Guan,W. and Condic,M.L.

TITLE Characterization of Netrin-1, Neogenin and cUNC-5H3 expression
 during chick dorsal root ganglia development

JOURNAL Gene Expr. Patterns 3, 369-373 (2003)

REFERENCE 2 (bases 1 to 2962)

AUTHORS Guan,W. and Condic,M.L.

TITLE Direct Submission

JOURNAL Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah,
 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA

FEATURES Location/Qualifiers

source

1. .2962
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"

CDS

1. .2796
 /codon_start=1
 /product="UNC5-like protein 3"
 /protein_id="AA067275.1"
 /db_xref="GI:31442351"

```
/translation="MGKGLEGTAARCGLMGYLLHSVLPALAVLGASRPGSAAQDDDF
FFHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPNVLYCKASPATQIYFKCNSEVHQ
KDHVVDERVEDETSGLIVCEVSI EISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPEGIPVAEVEWLKNEEVIDPVEDRNFYI
TIDHNLI IKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRC
GRGFQKRTRCTNPAPLNGGAFCEGQNVQKIACITTLCPVDGKWTSSWSKWSTCGTECTH
WRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPSDDDVALYVGIVIAVIVCLA
ISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
GPVYALHDVSDKIPMTNSPILDPPLNLKIKVYNTSGAVTPQDELSDFSSKLSPQITQS
LLENETLNVKNQSLARQTDPSCTAFGTFNLSLGGHLVIPNSGVSLLI PAGAVPQGRVYE
MYVTVHRKEGMRPVEDSQTLTTPVVS CGPPGALLTRPVVLTMHHC AEPNMDDWQIQIQL
KHQAGQGPPWEDVVVGEENFTTPCYIQLDPEACHILTETLSTYALVGQSITKAAAKRL
KLAI FGPLSCSSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEPKTLHFKGSTHN
LRLSIHDIAHSLWKSCLKPAKYQEIPFYHIWSGCQRNLHCTFTLERFSLNTLELVCKLC
VRQEVEGQGI FQLNCSVSEPTGIDYPIMDSAGSITTIVGPNAFSIFLPIRQKLCSSL
DAPQETRGRHDWRMLAHKCLKLDRLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
EMGRHETVVS LA AEGNY"
```

Query Match 36.0%; Score 991; DB 5; Length 2962;
Best Local Similarity 62.6%; Pred. No. 2.6e-156;
Matches 1662; Conservative 0; Mismatches 930; Indels 63; Gaps 5;

Db	638	ACCGAAATTTTACATCACCATTGATCACAACTGATCATCAAGCAAGCCCGGCTTTCCG	697
Qy	689	ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGACGCGCTCCG	748
Db	698	ACACGGCTAACTACACCTGTGTTGCCAAAAACATTGTGGCCAAAAGGAAAAGCACGACAG	757
Qy	749	CTGCTGTCATCGTCTACGTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCA	808
Db	758	CAACTGTGATTGTCTATGTGAATGGAGGCTGGTCTACCTGGACCGAGTGGTCAGCGTGCA	817
Qy	809	GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	868
Db	818	ACAGCCGCTGTGGGAGAGGCTTCCAGAAGCGCACAAAGGACCTGCACTAACCCTGCCCCAC	877
Qy	869	TCAACGGGGGCGCTTTCTGTGAGGGGCGAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT	928
Db	878	TCAATGGGGGGGCTTCTGCGAGGGGCGAAAATGTTTCAGAAAATAGCTTGCACCACCCTGT	937
Qy	929	GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA	988
Db	938	GTCCAGTGGATGGCAAATGGACGTCCTGGAGCAAGTGGTCCACTTGTGGCACAGAGTGTA	997
Qy	989	CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC	1048
Db	998	CCCACTGGCGCCGGAGGGAGTGACAGCTCCGGCCCCGAAGAATGGAGGCAAGGACTGTG	1057
Qy	1049	AGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTG	1108
Db	1058	AGGGACTGGTGTGCTGCAGTCTAAGAACTGCACTGATGGGCTCTGCATGCAGGCTGCACCTG	1117
Qy	1109	GCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCC	1165
Db	1118	ACTCGGATGATGTTGCTCTCTACGTGGGGATTGTGCTGCTGTGATTGTGTGCCTGGCTA	1177
Qy	1166	TGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATG	1225
Db	1178	TTTCTGTGGTTGTGGCCCTGTTTGTCTATCGCAAGAACCACCGTGACTTTGAGTCAGATA	1237
Qy	1226	TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG	1285
Db	1238	TTATCGACTCATCGGCGCTAAATGGGGGATTTTCAGCCTGTTAACATCAAGGCTGCAAGAC	1297
Qy	1286	CAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACC	1345
Db	1298	AAGAC-----CTCTTGGCAGTGCCACCAGACCTCACTTCTGCTGCAGCCATGTACA	1348
Qy	1346	AGGGCAGTCTCTGTCCCCGGCAGGATG-----GGCCAGCCCCAAGTTCAGCTCACCAA	1400
Db	1349	GGGGCCCTGTGTATGCCTTGCATGATGTCTCTGATAAAATCCCAATGACCAATTCTCCGA	1408
Qy	1401	T----GGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTC	1456
Db	1409	TCCTGGACCCACTGCCCAATCTGAAGATTAAAGTTTATAACACCTCTGGAGCAGTCACCC	1468
Qy	1457	CCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC-----	1498
Db	1469	CCCAGGATGAACTCTCTGACTTCTCTCCCAAGCTGTCCCCACAGATTACCCAGTCTCTGT	1528

Qy 1499 -----AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA 1534
 Db 1529 TGGAGAATGAGACTCTGAACGTGAAGAACCAAAGCCTTGACGGCAAACAGACCCATCCT 1588
 Qy 1535 ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA 1594
 Db 1589 GCACTGCATTTGGGACCTTCAACTCGTTAGGGGGCCACCTAGTAATTCCTAATTCAGGAG 1648
 Qy 1595 TCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA 1654
 Db 1649 TGAGCTTGCTGATCCCAGCAGGGGCTGTTCCCCAAGGAAGAGTCTATGAAATGTATGTGA 1708
 Qy 1655 CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTC 1714
 Db 1709 CAGTCCACAGGAAGGAGGGCATGAGACCACCTGTAGAAGACAGCCAGACGCTGCTGACAC 1768
 Qy 1715 CCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGG 1774
 Db 1769 CAGTGGTGAGCTGTGGCCCACCAGGAGCGCTGCTGACCCGACCCGTTGTGCTGACCATGC 1828
 Qy 1775 ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG 1834
 Db 1829 ACCACTGTGCTGAGCCCAACATGGATGACTGGCAGATCCAGCTCAAGCACCAGGCAGGCC 1888
 Qy 1835 AGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACT 1894
 Db 1889 AGGGACCATGGGAGGATGTAGTGGTGGTGGGGAGGAAAACCTTCACCACTCCATGCTACA 1948
 Qy 1895 GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG 1954
 Db 1949 TCCAGCTGGACCCAGAGGCCTGTCATATCCTGACGGAGACCCTCAGCACGTACGCCTTGG 2008
 Qy 1955 TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG 2014
 Db 2009 TGGGACAATCCATCACCAAAGCAGCAGCCAAACGTCTCAAATTGGCCATCTTTGGACCAC 2068
 Qy 2015 TGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATG 2074
 Db 2069 TGTCTGTTCCTCACTGGAGTACAGCATCCGCGTCTACTGCCTCGATGACACACAGGATG 2128
 Qy 2075 CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCAC 2134
 Db 2129 CCCTGAAGGAGTCCCTCCAGCTTGAGCGGCAGATGGGTGGGCAGCTGTTGGAGGAACCCA 2188
 Qy 2135 GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCA 2194
 Db 2189 AAACCTTGCATTTTAAGGGAAGTACCCACAACCTGCGCTTATCCATTCATGACATTGCCC 2248
 Qy 2195 GCTCCCTGTGGAAGAGTAAGCTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCT 2254
 Db 2249 ACTCTCTCTGGAAGAGCAAACCTGCCGGCTAAATACCAGGAGATTCTTTTACCACATCT 2308
 Qy 2255 GGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCA 2314
 Db 2309 GGAGTGGGTGCCAGAGGAACCTTGCACTGCACCTTCACGCTGGAACGATTTCAGTCTCAATA 2368

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 14:42:51 ; Search time 1401 Seconds
(without alignments)
11628.216 Million cell updates/sec

Title: US-10-624-932-1
Perfect score: 2752
Sequence: 1 ccgcggggcccccgcgcccg.....tgagtgtgaggccggccag 2752

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					Description
Result	Query						
No.	Score	Match	Length	DB	ID		
1	2752	100.0	2752	6	ABK37922		Abk37922 cDNA enco
2	2752	100.0	2752	12	ADH71617		Adh71617 Human gen
3	2687.4	97.7	2697	6	ABK52891		Abk52891 Human net
4	2676.4	97.3	2881	6	ABK49422		Abk49422 DNA encod

5	2676.4	97.3	2881	12	ADH71609	Adh71609 Human gen
6	2674.8	97.2	2881	12	ADH71649	Adh71649 Human gen
7	2674.8	97.2	2881	12	ADH71635	Adh71635 Human gen
8	2674.8	97.2	2881	12	ADH71637	Adh71637 Human gen
9	2674.8	97.2	2881	12	ADH71641	Adh71641 Human gen
10	2674.8	97.2	2881	12	ADH71629	Adh71629 Human gen
11	2674.8	97.2	2881	12	ADH71631	Adh71631 Human gen
12	2674.8	97.2	2881	12	ADH71645	Adh71645 Human gen
13	2674.8	97.2	2881	12	ADH71627	Adh71627 Human gen
14	2674.8	97.2	2881	12	ADH71639	Adh71639 Human gen
15	2674.8	97.2	2881	12	ADH71643	Adh71643 Human gen
16	2674.8	97.2	2881	12	ADH71625	Adh71625 Human gen
17	2674.8	97.2	2881	12	ADH71647	Adh71647 Human gen
18	2673.2	97.1	2881	10	ADG42568	Adg42568 Novel hum
19	2662.8	96.8	2880	12	ADH71633	Adh71633 Human gen
20	2368.6	86.1	2907	4	AAK52261	Aak52261 Human pol
21	2356	85.6	3561	12	ADL06497	Adl06497 Human tum
22	2356	85.6	3580	6	ABK15169	Abk15169 Human REP
23	2328	84.6	2575	12	ADH71621	Adh71621 Human gen
24	2266.2	82.3	2463	12	ADH71623	Adh71623 Human gen
25	2259	82.1	3014	2	AAV52940	Aav52940 Rat UNC-5
26	2252.2	81.8	2697	6	AAS16843	Aas16843 Rat netri
27	1635.4	59.4	2635	11	ADN95100	Adn95100 Human LEC
28	1562.4	56.8	1787	2	AAV52941	Aav52941 Human UNC
29	1304.2	47.4	9700	4	AAC90958	Aac90958 Plasmid p
30	1206.6	43.8	1321	4	AAH99530	Aah99530 Human pro
31	987.6	35.9	1002	12	ADH71619	Adh71619 Human gen
32	985.6	35.8	1009	12	ADH71613	Adh71613 Human gen
33	970	35.2	3646	5	AAS75738	Aas75738 DNA encod
34	970	35.2	3646	13	ADR99257	Adr99257 Human unc
35	968.8	35.2	2796	10	AAL56266	Aal56266 Human thr
36	943.4	34.3	993	12	ADH71611	Adh71611 Human gen
37	936.2	34.0	2860	6	ABT06279	Abt06279 Human NOV
38	936.2	34.0	2860	6	ABT06280	Abt06280 Human NOV
39	913.6	33.2	2895	6	ABQ93898	Abq93898 Human tra
40	905.6	32.9	3933	13	ADR07892	Adr07892 Full leng
41	904.2	32.9	2986	13	ADR99249	Adr99249 Human lRO
42	902.4	32.8	3884	4	AAS21316	Aas21316 Human cDN
43	902.4	32.8	3884	8	ACA03675	Aca03675 cDNA enco
44	902.4	32.8	3884	8	ABX89213	Abx89213 DNA encod
45	902.4	32.8	3884	8	ACD41867	Acd41867 Human sec

ALIGNMENTS

RESULT 1

ABK37922

ID ABK37922 standard; cDNA; 2752 BP.

XX

AC ABK37922;

XX

DT 21-MAY-2002 (first entry)

XX

DE cDNA encoding Human protein NOV1.

XX

KW Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;

KW cell signal processing disorder; metabolic disorder; obesity; infection;
 KW anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; immune disorder;
 KW haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
 KW osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
 KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
 KW psychosis; neurological disorder; anxiety; schizophrenia;
 KW manic depression; dementia; dyskinesia; Huntington's disease;
 KW Gilles de la Tourette's syndrome; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200210216-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-US024225.
 XX
 PR 28-JUL-2000; 2000US-0221409P.
 PR 04-AUG-2000; 2000US-0222840P.
 PR 04-AUG-2000; 2000US-0223752P.
 PR 04-AUG-2000; 2000US-0223762P.
 PR 04-AUG-2000; 2000US-0223769P.
 PR 04-AUG-2000; 2000US-0223770P.
 PR 14-AUG-2000; 2000US-0225146P.
 PR 15-AUG-2000; 2000US-0225392P.
 PR 15-AUG-2000; 2000US-0225470P.
 PR 16-AUG-2000; 2000US-0225697P.
 PR 01-FEB-2001; 2001US-0263662P.
 PR 05-APR-2001; 2001US-0281645P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigar M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
 PI Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;
 XX
 DR WPI; 2002-180074/23.
 DR P-PSDB; AAU85403.
 XX
 PT New isolated cytoplasmic, nuclear, membrane bound, or secreted
 PT polypeptide, useful for treating cardiomyopathy, atherosclerosis,
 PT infections, cancer, neurodegenerative, metabolic, hematopoietic and
 PT immune disorders.
 XX
 PS Claim 9; Page 9-10; 213pp; English.
 XX
 CC The invention relates to an isolated cytoplasmic, nuclear, membrane
 CC bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
 CC form. Also included are the nucleic acids encoding the NOVX proteins, a
 CC vector comprising the nucleic acid, a cell comprising the vector, an anti
 CC -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
 CC antibody are useful for treating or preventing a NOVX-associated
 CC disorder, where the disorder is selected from cardiomyopathy,
 CC atherosclerosis, diabetes, a disorder related to cell signal processing
 CC and metabolic pathway modulation, metabolic disorders, obesity,
 CC infectious disease, anorexia, cancer-associated cachexia, cancer,
 CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,

CC immune disorders, haematopoietic disorders, and the various
 CC dyslipidaemias, metabolic disturbances associated with obesity, the
 CC metabolic syndrome X and wasting disorders associated with chronic
 CC diseases, bacterial, fungal, protozoal and viral infections, pain,
 CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
 CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
 CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic
 CC hypertrophy, and psychotic and neurological disorders, including anxiety,
 CC schizophrenia, manic depression, delirium, dementia, and dyskinesias,
 CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
 CC nucleic acid is useful in gene therapy. The present sequence encodes a
 CC NOVX protein

XX

SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Query Match 100.0%; Score 2752; DB 6; Length 2752;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	60
Db	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	60
Qy	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Db	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Qy	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Qy	181	TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	181	TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC	360
Db	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC	360
Qy	361	ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	361	ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600

Db	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Qy	661	CTGGTGGTGCGACAGGCCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	661	CTGGTGGTGCGACAGGCCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Qy	721	ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Db	721	ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Qy	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCCGTGGAGC	960
Db	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCCGTGGAGC	960
Qy	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Db	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Qy	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Db	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Qy	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Db	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Qy	1141	ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Db	1141	ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Qy	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG	1260
Db	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG	1260
Qy	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC	1320
Db	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC	1320
Qy	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Db	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Qy	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440
Db	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440

Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Db	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Qy	1501	AAC TACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Db	1501	AAC TACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Qy	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Db	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Qy	1621	ATACCCCAGGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Db	1621	ATACCCCAGGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Qy	1681	TTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC	1740
Db	1681	TTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC	1740
Qy	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Db	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Qy	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Db	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Qy	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Db	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Qy	1921	GTCTTCACCGAGCAGCTGGGCGCCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980
Db	1921	GTCTTCACCGAGCAGCTGGGCGCCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980
Qy	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Db	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Qy	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Db	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Qy	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC	2160
Db	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC	2160
Qy	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Db	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Qy	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC	2280
Db	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC	2280

Qy	2281	TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG	2340
Db	2281	TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG	2340
Qy	2341	GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC	2400
Db	2341	GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC	2400
Qy	2401	ACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC	2460
Db	2401	ACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC	2460
Qy	2461	CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA	2520
Db	2461	CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA	2520
Qy	2521	CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCAT	2580
Db	2521	CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCAT	2580
Qy	2581	CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG	2640
Db	2581	CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG	2640
Qy	2641	CGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG	2700
Db	2641	CGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG	2700
Qy	2701	CCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2701	CCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG	2752

RESULT 2

ADH71617

ID ADH71617 standard; DNA; 2752 BP.

XX

AC ADH71617;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21e SEQ ID NO:513.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.

PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-0423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71618.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 513; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Query Match 100.0%; Score 2752; DB 12; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGGCCATGGCCGTCCGGCCC	60
Db	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGGCCATGGCCGTCCGGCCC	60
Qy	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Db	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Qy	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Qy	181	TTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	181	TTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Db	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Qy	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Qy	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTACGTGAACGGTGGGTGG	780
Db	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTACGTGAACGGTGGGTGG	780

Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Qy	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC	960
Db	901	GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC	960
Qy	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Db	961	AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA	1020
Qy	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Db	1021	GCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC	1080
Qy	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Db	1081	AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC	1140
Qy	1141	ATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Db	1141	ATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTGCTTGTCTCATCCTCGTTTATTGCCGG	1200
Qy	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG	1260
Db	1201	AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG	1260
Qy	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC	1320
Db	1261	CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC	1320
Qy	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Db	1321	CTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC	1380
Qy	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440
Db	1381	CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1440
Qy	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Db	1441	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1500
Qy	1501	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Db	1501	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC	1560
Qy	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620
Db	1561	CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC	1620

Qy	1621	ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Db	1621	ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG	1680
Qy	1681	TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC	1740
Db	1681	TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC	1740
Qy	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Db	1741	GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC	1800
Qy	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Db	1801	AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC	1860
Qy	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Db	1861	CTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC	1920
Qy	1921	GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980
Db	1921	GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC	1980
Qy	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Db	1981	GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAAC	2040
Qy	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Db	2041	ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG	2100
Qy	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC	2160
Db	2101	AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC	2160
Qy	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Db	2161	CACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTT	2220
Qy	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC	2280
Db	2221	GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC	2280
Qy	2281	TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG	2340
Db	2281	TGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG	2340
Qy	2341	GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAAGCATCAACTTCAACATCACCAAGGAC	2400
Db	2341	GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAAGCATCAACTTCAACATCACCAAGGAC	2400
Qy	2401	ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTCCCAGCCCTGGTGGGC	2460
Db	2401	ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTCCCAGCCCTGGTGGGC	2460
Qy	2461	CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA	2520

```

      |||
Db      2461 CCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
Qy      2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCAT 2580
      |||
Db      2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGACAGCCAT 2580
Qy      2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
      |||
Db      2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Qy      2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
      |||
Db      2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Qy      2701 CCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
      |||
Db      2701 CCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2752

```

RESULT 3

ABK52891

ID ABK52891 standard; DNA; 2697 BP.

XX

AC ABK52891;

XX

DT 27-AUG-2002 (first entry)

XX

DE Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.

XX

KW Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;

KW nootropic; neuroprotective; cytostatic; antiparkinsonian;

KW cerebroprotective; cancer; central nervous system; CNS; stroke;

KW Parkinson's disease; multiple sclerosis; Alzheimer's disease.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	1. .2697
----	-----	----------

FT		/*tag= a
----	--	----------

FT		/product= "Netrin binding membrane receptor UNC5H-1"
----	--	--

XX

PN WO200233080-A2.

XX

PD 25-APR-2002.

XX

PF 15-OCT-2001; 2001WO-EP011891.

XX

PR 16-OCT-2000; 2000US-0240061P.

XX

PA (FARB) BAYER AG.

XX

PI Koehler RH;

XX

DR WPI; 2002-463314/49.

DR P-PSDB; AAU97899.

XX

PT Novel human netrin binding membrane receptor polypeptide and
PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease.

XX

PS Claim 1; Fig 1; 94pp; English.

XX

CC This invention relates to the DNA and protein sequences of a novel
CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC sequence of the invention is useful as a probe for detecting a nucleic
CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC of the invention are useful to screen for agents which decrease the
CC activity of the UNC5H-1 protein. The sequences are also useful for
CC screening agents which regulate (modulate) the activity of the protein of
CC the invention. A pharmaceutical composition containing the protein of the
CC invention or a reagent that modulates the activity of the UNC5H-1 protein
CC may be useful for treating a UNC5H-1 dysfunction related disease such as
CC cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC proteins comprising the UNC5H-1 protein are useful for generating
CC antibodies and for in various assay systems, and the protein can be used
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC of the invention is useful for detecting a coding sequence for the UNC5H-
CC 1 protein. The present sequence represents a DNA sequence encoding the
CC human netrin binding membrane receptor UNC5H-1 protein of the invention

XX

SQ Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Query Match 97.7%; Score 2687.4; DB 6; Length 2697;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 46 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
|
Db 1 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy 106 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
|
Db 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy 166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
|
Db 121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy 226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
|
Db 181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy 286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345
|
Db 241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy 346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGGGCTGGAG 405
|
Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGGGCTGGAG 360

Qy 406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Db	361		GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Qy	466		TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	525
Db	421		TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Qy	526		TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	585
Db	481		TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Qy	586		GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	645
Db	541		GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Qy	646		ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACCTACACC	705
Db	601		ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACCTACACC	660
Qy	706		TGCGTGGCCAAGAATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCATCGTCTAC	765
Db	661		TGCGTGGCCAAGAATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Qy	766		GTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	825
Db	721		GTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Qy	826		GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	885
Db	781		GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Qy	886		TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	945
Db	841		TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Qy	946		TGGAGCCCGTGGAGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Db	901		TGGAGCCCGTGGAGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Qy	1006		GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Db	961		GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Qy	1066		ACCCGCAACTGTACAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
Db	1021		ACCCGCAACTGTACAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Qy	1126		CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCCTCATC	1185
Db	1081		CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCCTCATC	1140
Qy	1186		CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Db	1141		CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Qy	1246		ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305

Db	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Qy	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Db	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
Qy	1366	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
Db	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Qy	1426	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485
Db	1381	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Qy	1486	CGCCTCTCCACCCAGAATACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1441	CGCCTCTCCACCCAGAATACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC	1560
Qy	1606	ATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1561	ATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Qy	1666	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Qy	1726	TGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1681	TGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Qy	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1906	GCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1861	GCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1921	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2145
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2100

Qy	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205
Db	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Qy	2206	AAGAGTAAGCTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2265
Db	2161	AAGAGTAAGCTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Qy	2266	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG	2325
Db	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTG	2280
Qy	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
Db	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Qy	2386	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2445
Db	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2400
Qy	2446	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGCAGAAGATAATT	2505
Db	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGCAGAAGATAATT	2460
Qy	2506	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTC	2565
Db	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTC	2520
Qy	2566	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2625
Db	2521	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Qy	2626	AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2685
Db	2581	AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2640
Qy	2686	GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGA	2742
Db	2641	GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGA	2697

RESULT 4

ABK49422

ID ABK49422 standard; DNA; 2881 BP.

XX

AC ABK49422;

XX

DT 15-JUL-2002 (first entry)

XX

DE DNA encoding human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;

KW cell signal processing; metabolic pathway modulation; cancerous tissue;

KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;

KW chromosome 13; gene; ds.

XX

OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 87. .2786
 FT /*tag= a
 FT /product= "Human UNC5-like protein NOV1"
 XX
 PN WO200229038-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-US031377.
 XX
 PR 04-OCT-2000; 2000US-0237862P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Herrmann JL, Rastelli L, Shimkets RA;
 XX
 DR WPI; 2002-340104/37.
 DR P-PSDB; AAU79939.
 XX
 PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
 PT treating cardiomyopathy, arteriosclerosis, and cancer.
 XX
 PS Claim 8; Page 7-8; 180pp; English.
 XX
 CC The present invention relates to a new NOVX polypeptide having a 900
 CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
 CC residue amino acid sequence, as given in the specification. The novel
 CC polypeptide, and its encoding polynucleotide, are used to treat
 CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
 CC signal processing and metabolic pathway modulation, in a human. Detecting
 CC the polypeptide or polynucleotide is useful for identifying cancerous
 CC tissue. The antibody can be used to treat diabetes or cancer. The host
 CC cells can be used to produce non-human transgenic animals useful in drug
 CC screening. The present nucleic acid sequence is that of the human UNC5-
 CC like NOV1 gene located on chromosome 13. This sequence encodes the human
 CC UNC5-like protein NOV1 of the invention
 XX
 SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 97.3%; Score 2676.4; DB 6; Length 2881;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

Qy 1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101
 Qy 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
 Qy 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy	181	TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	222	TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	281
Qy	241	AAGGCCGTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	282	AAGGCCGTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Db	342	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC	401
Qy	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	402	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061

Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914

Db	1899		CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915		TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959		TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975		GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019		GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035		TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079		TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095		CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139		CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155		AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199		AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215		CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259		CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275		TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319		TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335		CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379		CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395		AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439		AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455		GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499		GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515		GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559		GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575		AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619		AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635		GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679		GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695		GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2752

RESULT 5

ADH71609

ID ADH71609 standard; DNA; 2881 BP.

XX

AC ADH71609;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21a SEQ ID NO:505.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;

Qy	361	ATTAATGTCTCAAGGCAGCAGGTCGAGAAAGGTGTTTCGGGCTGGAGGAATACTGGTGCCAG	420
Db	402	ATTAATGTCTCAAGGCAGCAGGTCGAGAAAGGTGTTTCGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTCGCCGACGCGCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTCGCCGACGCGCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCCTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCCTGG	1001
Qy	958	AGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257

Db	1242	 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	 AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	 ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	 AGGTTGCCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	 CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094

Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG	2796

RESULT 6

ADH71649

ID ADH71649 standard; DNA; 2881 BP.

XX

AC ADH71649;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21u SEQ ID NO:545.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.

PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71650.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 545; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,

CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
 CC antilipaemic activity, and may have a use in gene therapy, and as a
 CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
 CC any of the 303 fully defined nucleotide sequences given in the
 CC specification. The polypeptide is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease. The
 CC polypeptide, polynucleotide and antibody are useful in diagnosing,
 CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
 CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
 CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
 CC further used as hybridisation probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. The present sequence
 CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy	1	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	60
Db	42	CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	101
Qy	61	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Db	102	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	161
Qy	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	162	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	221
Qy	181	TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	222	TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	281
Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	282	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Db	342	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC	401
Qy	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	402	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600

Db	582	 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	780
Db	762	 ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1197
Db	1182	 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437

Db	1419	AGCCCCAAGTTCAGCTACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318

Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTGGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCTCATTGGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGCGG	2678
Qy	2635	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG	2796

RESULT 7

ADH71635

ID ADH71635 standard; DNA; 2881 BP.

XX

AC ADH71635;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21n SEQ ID NO:531.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.

PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71636.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 531; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy      1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGGCCATGGCCGTCCGGCCC 60
          |||
Db      42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGGCCATGGCCGTCCGGCCC 101

Qy     61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
          |||
Db    102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161

Qy    121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
          |||
Db    162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy    181 TTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
          |||
Db    222 TTCCTGGTGGAGCCCAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281

Qy    241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
          |||
Db    282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341

Qy    301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC 360
          |||
Db    342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC 401

Qy    361 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGGCTGGAGGAATACTGGTGCCAG 420
          |||
Db    402 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGGCTGGAGGAATACTGGTGCCAG 461

Qy    421 TGGCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
          |||
Db    462 TGGCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521

Qy    481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
          |||
Db    522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581

Qy    541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
          |||
Db    582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641

Qy    601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660
          |||
Db    642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701

Qy    661 CTGGTGGTGGCAGAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720
          |||
Db    702 CTGGTGGTGGCAGAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761

Qy    721 ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 780
          |||
```

Db	762	ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	AACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	AACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658

Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTCTAGACCCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTCTAGACCCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACATCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498

Qy 2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTG 2514
 |||
 Db 2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTG 2558
 Qy 2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC 2574
 |||
 Db 2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC 2618
 Qy 2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634
 |||
 Db 2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
 Qy 2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694
 |||
 Db 2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
 Qy 2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
 |||
 Db 2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2796

RESULT 8

ADH71637

ID ADH71637 standard; DNA; 2881 BP.

XX

AC ADH71637;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21o SEQ ID NO:533.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.

PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71638.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 533; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60
|
Db 42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101

Qy 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
|

Db	102	GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGGCTCCGCGGCTCGGGTGCC	161
Qy	121	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	162	CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	221
Qy	181	TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGTCTGTGTGC	240
Db	222	TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGTCTGTGTGC	281
Qy	241	AAGGCCGTGCCC GCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	282	AAGGCCGTGCCC GCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Db	342	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC	401
Qy	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGGCTGGAGGAATACTGGTGCCAG	420
Db	402	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001

Qy	958	AGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838

Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCACCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635	GAGGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694

```

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
Qy      2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2796

```

RESULT 9

ADH71641

ID ADH71641 standard; DNA; 2881 BP.

XX

AC ADH71641;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21q SEQ ID NO:537.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;

PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71642.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

PS Example 21; SEQ ID NO 537; 1880pp; English.

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

SQ Sequence 2881 BP; 526 A; 986 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 CCGCGGGGGCCCCGCGCCCGGGCCCGCCGCGCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCC 60

Db 42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101

Qy 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
|||

Db 102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161

Qy 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180

Db 162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240

Db 222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281

Qy 241 AAGGCCGTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300

Db 282 AAGGCCGTGCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341

Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Db	342	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC	401
Qy	361	ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTTCGGGCTGGAGGAATACTGGTGCCAG	420
Db	402	ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTTCGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181

Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1241
Qy	1198	CGBAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGBAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTCACCATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCCAGCTCACCATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACC CGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACC CGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034

Db	2019		GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035		TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079		TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGCTGGTGCAG	2138
Qy	2095		CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139		CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155		AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199		AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215		CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259		CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275		TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319		TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335		CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379		CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395		AAGGACACAAGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439		AAGGACACAAGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455		GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499		GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTTCGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515		GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559		GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575		AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619		AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635		GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679		GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695		GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739		GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2796

RESULT 10

ADH71629

ID ADH71629 standard; DNA; 2881 BP.

XX

AC ADH71629;

XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NOV21k SEQ ID NO:525.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.

PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-0423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71630.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGGCAGAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGGCAGAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCCTACCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCCTACCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377

Db	1362		GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378		AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419		AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438		ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479		ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498		CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539		CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558		TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599		TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618		GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659		GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678		AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719		AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738		GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779		GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798		GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839		GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855		CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899		CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915		TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959		TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975		GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019		GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035		TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079		TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095		CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139		CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155		AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2214

Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTCTAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTCTAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG	2796

RESULT 11

ADH71631

ID ADH71631 standard; DNA; 2881 BP.

XX

AC ADH71631;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV211 SEQ ID NO:527.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71632.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 527; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,

CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 984 C; 868 G; 503 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy      1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60
          |||
Db      42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101

Qy      61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGGTGCC 120
          |||
Db     102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGGTGCC 161

Qy     121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
          |||
Db     162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy     181 TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
          |||
Db     222 TTCCTGGTGGAGCCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281

Qy     241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
          |||
Db     282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341

Qy     301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
          |||
Db     342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC 401

Qy     361 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGGCTGGAGGAATACTGGTGCCAG 420
          |||
Db     402 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGCGGGCTGGAGGAATACTGGTGCCAG 461

Qy     421 TCGGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
          |||
Db     462 TCGGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521

Qy     481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
          |||
Db     522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581

Qy     541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
          |||
Db     582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641

Qy     601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660
          |||
Db     642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701

Qy     661 CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720
```

Db	702	 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTTCATCGTCTACGTGAACGGTGGGTGG	780
Db	762	 ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTTCATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	 AGCAAGTGGTTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1197
Db	1182	 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	 CGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	 ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557

Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GTCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAAGCATCAACTTCAACATCACC	2438

Qy 2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
 |||
 Db 2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
 Qy 2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGCAGAGAATAATTTCCAGCCTG 2514
 |||
 Db 2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGCAGAGAATAATTTCCAGCCTG 2558
 Qy 2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC 2574
 |||
 Db 2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC 2618
 Qy 2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634
 |||
 Db 2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
 Qy 2635 GAGGCGCGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694
 |||
 Db 2679 GAGGCGCGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
 Qy 2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
 |||
 Db 2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCTGGAGGCTGAGTGCTGAGGCCGGCCAG 2796

RESULT 12

ADH71645

ID ADH71645 standard; DNA; 2881 BP.

XX

AC ADH71645;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21s SEQ ID NO:541.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71646.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 541; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60

Db	42		CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC	101
Qy	61		GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	120
Db	102		GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC	161
Qy	121		CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	180
Db	162		CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC	221
Qy	181		TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	240
Db	222		TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	281
Qy	241		AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	282		AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Qy	301		GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC	360
Db	342		GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC	401
Qy	361		ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	402		ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421		TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462		TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481		AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522		AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541		ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582		ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601		AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642		AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661		CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702		CTGGTGGTGGGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721		ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	780
Db	762		ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTATCGTCTACGTGAACGGTGGGTGG	821
Qy	781		TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822		TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841		AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900

Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCCTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAAGCATCAAGCCAGCAAAGCAGACAACCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAAGCATCAAGCCAGCAAAGCAGACAACCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	AACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	AACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT	1737
Db	1719	AGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT	1778

Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGCGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGCGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618

Qy 2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634
 |||
 Db 2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
 Qy 2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694
 |||
 Db 2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
 Qy 2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
 |||
 Db 2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2796

RESULT 13

ADH71627

ID ADH71627 standard; DNA; 2881 BP.

XX

AC ADH71627;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21j SEQ ID NO:523.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P..
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX

Db	222	TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC	281
Qy	241	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
Db	282	AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Qy	301	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC	360
Db	342	GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC	401
Qy	361	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	420
Db	402	ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG	461
Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTGCGCGCAGCGCCTCCGCTGCTGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121

Qy	1078	ACCAAGTACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAAGTACCTCTGTGTGCACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTGAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958

Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGAGGCTGAGTGCTGAGGCCGGCCAG	2796

RESULT 14

ADH71639

ID ADH71639 standard; DNA; 2881 BP.

XX

AC ADH71639;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21p SEQ ID NO:535.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389120P.

PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71640.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 535; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy      1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60
          |||
Db      42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101

Qy      61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
          |||
Db     102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161

Qy     121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
          |||
Db     162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy     181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
          |||
Db     222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281

Qy     241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
          |||
Db     282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341

Qy     301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGTCCGC 360
          |||
Db     342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC 401

Qy     361 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 420
          |||
Db     402 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 461
```

Qy	421	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	480
Db	462	TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC	521
Qy	481	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	540
Db	522	AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC	581
Qy	541	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	600
Db	582	ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG	641
Qy	601	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTCATCGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTGCGCCGAGCGCCTCCGCTGCTGTCTCATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCATCCTCGTTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301

Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	AACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497
Db	1479	AACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCAGCCCT	1797
Db	1779	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	GCTGCCGCCAAGCGCCTCAGGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGAC	2154

Db	2139	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	CTCCTTGTGCTAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334
Db	2319	TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2378
Qy	2335	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2394
Db	2379	CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC	2438
Qy	2395	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2454
Db	2439	AAGGACACAAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG	2498
Qy	2455	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGGCAGAAGATAATTTCCAGCCTG	2514
Db	2499	GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGGGCAGAAGATAATTTCCAGCCTG	2558
Qy	2515	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2574
Db	2559	GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC	2618
Qy	2575	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2634
Db	2619	AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG	2678
Qy	2635	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2694
Db	2679	GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG	2738
Qy	2695	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG	2752
Db	2739	GGCCAGCCAGACGCTGGCCTCTTCACAGTGTGCGGAGGCTGAGTGCTGAGGCCGGCCAG	2796

RESULT 15

ADH71643

ID ADH71643 standard; DNA; 2881 BP.

XX

AC ADH71643;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21r SEQ ID NO:539.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71644.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 539; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising

CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy      1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60
          |||
Db      42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101

Qy      61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
          |||
Db     102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161

Qy     121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
          |||
Db     162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221

Qy     181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
          |||
Db     222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281

Qy     241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
          |||
Db     282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341

Qy     301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
          |||
Db     342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGGAGCCGACCATGGAGGTCCGC 401

Qy     361 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 420
          |||
Db     402 ATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTCTGGGCTGGAGGAATACTGGTGCCAG 461

Qy     421 TCGGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
          |||
Db     462 TCGGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521

Qy     481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
          |||
Db     522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581

Qy     541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
          |||
Db     582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
```

Qy	601	AACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	660
Db	642	AACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC	701
Qy	661	CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACCTACACCTGCGTGGCCAAGAAC	720
Db	702	CTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACCTACACCTGCGTGGCCAAGAAC	761
Qy	721	ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	780
Db	762	ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG	821
Qy	781	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	840
Db	822	TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG	881
Qy	841	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	900
Db	882	AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT	941
Qy	901	GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCAGTAGACGGCAGCTGGAGCCCGTGG	957
Db	942	GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG	1001
Qy	958	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1017
Db	1002	AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC	1061
Qy	1018	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1077
Db	1062	CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT	1121
Qy	1078	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1137
Db	1122	ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC	1181
Qy	1138	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCTATCCTCGTTTATTGC	1197
Db	1182	CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCTCTATCCTCGTTTATTGC	1241
Qy	1198	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1257
Db	1242	CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1301
Qy	1258	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1317
Db	1302	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG	1361
Qy	1318	GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1377
Db	1362	GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC	1418
Qy	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1437
Db	1419	AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC	1478
Qy	1438	ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497

Db	1479	 ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538
Qy	1498	CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1557
Db	1539	 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC	1598
Qy	1558	TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1617
Db	1599	 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT	1658
Qy	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1677
Db	1659	 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG	1718
Qy	1678	AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1737
Db	1719	 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT	1778
Qy	1738	GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1797
Db	1779	 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT	1838
Qy	1798	GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG	1854
Db	1839	 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG	1898
Qy	1855	CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Db	1899	 CTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	1958
Qy	1915	TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Db	1959	 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Qy	1975	GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2034
Db	2019	 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG	2078
Qy	2035	TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2094
Db	2079	 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG	2138
Qy	2095	CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2154
Db	2139	 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC	2198
Qy	2155	AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2214
Db	2199	 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAG	2258
Qy	2215	CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2274
Db	2259	 CTCCTTGTGAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC	2318
Qy	2275	TTGCACTGCACCTTACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG	2334

Db 2319 TTGCACTGCACCTTCACCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
 Qy 2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2438
 Qy 2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
 Qy 2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGCGCAGAAGATAATTTCCAGCCTG 2514
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTGCGCAGAAGATAATTTCCAGCCTG 2558
 Qy 2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC 2574
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAACTCCACCTGGAC 2618
 Qy 2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
 Qy 2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
 Qy 2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2796

Search completed: March 5, 2005, 22:58:50
 Job time : 1411 secs

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 19:26:23 ; Search time 433.33 Seconds
(without alignments)
10391.692 Million cell updates/sec

Title: US-10-624-932-1
Perfect score: 2752
Sequence: 1 ccgcggggcccccgcgcccg.....tgagtgctgaggccggccag 2752

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2259	82.1	3014	2	US-08-808-982-1	Sequence 1, Appli
2	2259	82.1	3014	3	US-09-306-902A-1	Sequence 1, Appli
3	1562.4	56.8	1787	2	US-08-808-982-2	Sequence 2, Appli
4	1562.4	56.8	1787	3	US-09-306-902A-2	Sequence 2, Appli
5	841.4	30.6	2831	2	US-08-808-982-3	Sequence 3, Appli
6	841.4	30.6	2831	3	US-09-306-902A-3	Sequence 3, Appli
7	833.6	30.3	3008	4	US-09-949-016-4794	Sequence 4794, Ap
8	487	17.7	2736	4	US-09-969-532-9	Sequence 9, Appli
9	487	17.7	3411	4	US-09-969-532-33	Sequence 33, Appl
10	467	17.0	2703	4	US-09-969-532-11	Sequence 11, Appl
11	460.2	16.7	2694	4	US-09-969-532-13	Sequence 13, Appl

12	439.8	16.0	2661	4	US-09-969-532-15	Sequence 15, Appl
13	323	11.7	349	4	US-09-471-276-345	Sequence 345, App
14	293.2	10.7	1968	4	US-09-969-532-31	Sequence 31, Appl
15	293.2	10.7	2001	4	US-09-969-532-29	Sequence 29, Appl
16	293.2	10.7	2010	4	US-09-969-532-27	Sequence 27, Appl
17	293.2	10.7	2043	4	US-09-969-532-25	Sequence 25, Appl
18	284.8	10.3	1659	4	US-09-969-532-7	Sequence 7, Appli
19	284.8	10.3	1692	4	US-09-969-532-5	Sequence 5, Appli
20	284.8	10.3	1701	4	US-09-969-532-3	Sequence 3, Appli
21	284.8	10.3	1734	4	US-09-969-532-1	Sequence 1, Appli
c 22	269	9.8	771	1	US-08-253-155A-17	Sequence 17, Appl
c 23	182.2	6.6	19326	4	US-09-949-016-16776	Sequence 16776, A
24	123.4	4.5	114139	4	US-09-949-016-16536	Sequence 16536, A
25	119	4.3	305	2	US-08-808-982-4	Sequence 4, Appli
26	119	4.3	305	3	US-09-306-902A-4	Sequence 4, Appli
c 27	80.4	2.9	601	4	US-09-949-016-169999	Sequence 169999,
28	80.2	2.9	966	4	US-09-969-532-23	Sequence 23, Appl
29	75.8	2.8	1008	4	US-09-969-532-19	Sequence 19, Appl
c 30	73.8	2.7	601	4	US-09-949-016-170033	Sequence 170033,
c 31	66.4	2.4	601	4	US-09-949-016-170040	Sequence 170040,
c 32	66.4	2.4	601	4	US-09-949-016-170041	Sequence 170041,
33	63.6	2.3	5784	4	US-09-949-016-462	Sequence 462, App
c 34	63	2.3	601	4	US-09-949-016-170000	Sequence 170000,
c 35	63	2.3	601	4	US-09-949-016-170034	Sequence 170034,
36	60.2	2.2	2820	4	US-09-854-845-15	Sequence 15, Appl
37	60.2	2.2	2865	4	US-09-854-845-13	Sequence 13, Appl
38	60.2	2.2	3105	4	US-09-854-845-5	Sequence 5, Appli
39	60.2	2.2	3150	4	US-09-854-845-1	Sequence 1, Appli
40	60.2	2.2	3237	4	US-09-854-845-7	Sequence 7, Appli
41	60.2	2.2	3282	4	US-09-854-845-3	Sequence 3, Appli
42	60.2	2.2	3411	4	US-09-854-845-11	Sequence 11, Appl
43	60.2	2.2	3456	4	US-09-854-845-9	Sequence 9, Appli
44	60.2	2.2	4074	4	US-09-854-845-17	Sequence 17, Appl
45	56.4	2.0	5178	4	US-09-949-016-5241	Sequence 5241, Ap

ALIGNMENTS

RESULT 1

US-08-808-982-1

; Sequence 1, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

```

;      ZIP: 94104
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/808,982
;      FILING DATE:
;      CLASSIFICATION: 530
;      ATTORNEY/AGENT INFORMATION:
;      NAME: OSMAN, RICHARD A
;      REGISTRATION NUMBER: 36,627
;      REFERENCE/DOCKET NUMBER: UC96-217
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 343-4341
;      TELEFAX: (415) 343-4342
;      INFORMATION FOR SEQ ID NO: 1:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 3014 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: cDNA
US-08-808-982-1

```

```

Query Match          82.1%; Score 2259; DB 2; Length 3014;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2427; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

```

```

Qy      46 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCCGTCCGGCCCGGCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy     106 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
        || || ||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGTGGTTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGGCCGGTGCCAACCCC 120

Qy     166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
        |||||||| |||||||||||||| |||| |||| |||||||| ||||||||||
Db     121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy     226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
        ||| || | |||||||||| |||| |||| ||||||||||||||||||||
Db     181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy     286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345
        ||||| |||||||| || ||||| || || |||||||| ||| | ||||| || |||
Db     241 TGGGTCCGCCAGGTGCATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy     346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAAGGTGTTTCGGGCTGGAG 405
        |||||||||||||| || || || || |||||||||| |||| |||| ||||||||
Db     301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy     406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465
        |||||||||||||| |||||||||||||||||| |||||||| ||||||||||
Db     361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

```

Qy	466	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	525
Db	421	TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG	480
Qy	526	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGAGGGGCATCCCTCCAGCCGAG	585
Db	481	TCCTGGAGCAAGGCATTGTACTACCTTGTGCCCCCAGAAGGAATCCCCCAGCTGAG	540
Qy	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	645
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
Qy	646	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCCTTGCTGACACGGCCAACTACACC	705
Db	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC	660
Qy	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGAGCGCCTCCGCTGCTGTCATCGTCTAC	765
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCAATTGTTTAT	720
Qy	766	GTGAACGGTGGGTGGTTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	825
Db	721	GTGAACGGTGGGTGGTTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT	780
Qy	826	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	885
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC	840
Qy	886	TGTGAGGGGCAGAAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC	945
Db	841	TGTGAGGGGCAGAAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCAGTGGATGGGAGC	900
Qy	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qy	1006	GAGTGTCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Db	961	GAGTGTCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC	1020
Qy	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qy	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCTGCTGCTTGTCTCTCATC	1185
Db	1081	CTCTACATCGGCCTTGTGCTGTGGCTGTGTGCTCTTCTTGCTGTTGCTGGCCCTTGA	1140
Qy	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Db	1141	CTCATTTACTGTGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC	1200
Qy	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Db	1201	ACCTCGGGCTTCCAGCCTGTGAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260

Qy	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Db	1261		
Qy	1366	ACCATCCAGCCAGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTATGTTTCGAGG	1320
Qy	1366	CAGGATGGGCCCAGCCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
Db	1321		
Qy	1366	CAGGATGGACCCAGCCCCAAGTTCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG	1380
Qy	1426	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485
Db	1381		
Qy	1426	AGTGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440
Qy	1486	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Db	1441		
Qy	1486	CGCCTCTCCACCCAAAATACTTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC	1500
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Db	1501		
Qy	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC	1560
Qy	1606	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Db	1561		
Qy	1606	ATACCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qy	1666	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725
Db	1621		
Qy	1666	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC	1680
Qy	1726	TGTGGACCCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Db	1681		
Qy	1726	TGTGGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA	1740
Qy	1786	GAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Db	1741		
Qy	1786	GAGCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGTTGG	1800
Qy	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Db	1801		
Qy	1846	GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1906	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1965
Db	1861		
Qy	1906	GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC	1920
Qy	1966	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
Db	1921		
Qy	2026	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG	1980
Qy	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Db	1981		
Qy	2026	TCCCTTGAGTACAACATCCGAGTGTAAGTGCCTACACGACACCCACGACGCTCTCAAGGAG	2040
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2145
Db	2041		
Qy	2086	GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC	2100
Qy	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205